Running head: END ANCHORING IN STM

End anchoring in short-term order memory

Simon Farrell and Anna Lelièvre

Department of Psychology

University of Bristol

Draft of September 17, 2008.

Dr. Simon Farrell

Department of Experimental Psychology

University of Bristol

12a Priory Road

Clifton, Bristol BS8 1TU

UK

e-mail: Simon.Farrell@bristol.ac.uk

fax: +44 (0)117 928 8588

Abstract

Temporally grouping lists has systematic effects on immediate serial recall accuracy, order errors, and recall latencies, and is generally taken to reflect the use of multiple dimensions of ordering in short-term memory. It has been argued that these representations are fully relative, in that all sequence positions are anchored to both the start and end of sequences. A comparison of four computational models of serial recall is presented that shows that the extant empirical evidence does not point towards fully relative positional markers, and is consistent with a simpler scheme in which only terminal items are coded with respect to the end of a sequence or subsequence. Results from the application of the models to data from two new experiments varying the size of groups in serially recalled lists support this conclusion.

Keywords: serial recall; short-term memory; temporal grouping; sequence memory; computational models; model selection; anchors.

End anchoring in short-term order memory

Evidence from investigations across psychology, including word recognition (e.g., Davis & Bowers, 2004), economic and valuative judgements (e.g., Hsee, Hastie, & Chen, 2008; Stewart, Chater, & Brown, 2006), spatial representation (Sadalla, Burroughs, & Staplin, 1980) and perception and absolute judgement (e.g., Bressan, 2006; Gravetter & Lockhead, 1973; Stewart, Brown, & Chater, 2005), have converged on the conclusion that in many cases the representation or judgement of objects and values is relative in nature. For example, although word recognition models have commonly assumed that letter information is encoded in an absolute, position specific way (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; McClelland & Rumelhart, 1981), recent evidence on letter confusions in briefly presented letter strings shows that transpositions of letters from different positions within words are more frequent than expected from absolute models, implying relativity in the way letters within words are represented (Davis & Bowers, 2004). In a similar vein, a striking phenomenon from a number of domains is that of the range effect, whereby participants' judgements of the physical properties of objects such as lightness (e.g., Bressan, 2006), length (e.g., Lacouture, 1997), and sweetness (Lawless, Horne, & Spiers, 2000) are shifted according to the range of stimuli to which the observer has been exposed.

One common assumption is that participants use some form of anchoring, whereby the representations or judgements of stimuli are anchored to landmarks including the extreme stimuli presented in an experiment (e.g. Braida et al., 1984; S. D. Brown, Marley, Donkin, & Heathcote, 2008). In this paper, we consider whether a similar principle of anchoring is at play in short-term order memory. Specifically, we ask to what extent the coding of position of items or events is anchored to both the start and end of that sequence (Henson, 1998b; Houghton, 1990). As we discuss below, this issue is of particular significance given one idiosyncratic characteristic of temporal order: the length of a sequence may often be unknown until the entire sequence has been presented, meaning that the end anchor is not usable as a stable referent during encoding. Below, we discuss a number of models of positional representation in short-term order memory, and present a series of simulations and experiments addressing the role of anchoring in short-term order memory.

Varieties of positional representations in short-term memory

A major focus of models of short-term memory has been the representations or mechanisms by which the order of sequences is remembered. One class of models, primacy models, assumes that items are represented by a gradient in the strength of activations or associations across a sequence, such that earlier items are more accessible (Farrell & Lewandowsky, 2002; Page & Norris, 1998). This simple ordinal scheme has been demonstrated to be sufficient to account for a number of benchmark data from the serial recall task, including serial position effects in accuracy and latency, the locality in positional confusions (nearby items tend to be confused), list length effects, and other phenomenon such as modality and word frequency effects (Farrell & Lewandowsky, 2002, 2004; Page & Norris, 1998). However, there exist data which are incompatible with the basic gradient-based representation assumed in these models.

One major objection to primacy models comes from the multifarious effects of grouping on serial recall performance. Grouping a list into subsequences by inserting pauses between groups (e.g., Maybery, Parmentier, & Jones, 2002; Ryan, 1969a, 1969b), intonation (Frankish, 1995; Reeves, Schmauder, & Morris, 2000), or simply suggesting a grouping structure through verbal instructions (Farrell, 2008; Wickelgren, 1964) lead to a number of well-replicated effects on recall. When examining accuracy, grouping leads to the appearance of "mini" serial position curves for groups, each group with its own primacy and recency. Grouping also has systematic effects on recall latencies: participants leave longer pauses in their output at group boundaries (Farrell, 2008; Maybery et al., 2002). Most problematic for primacy models is the effects of grouping on recall errors, particularly those involving confusions of items between groups (Henson, 1996). When adjacent positional confusions are examined, confusions between groups tend to dominate for ungrouped lists, while confusions within groups dominate for grouped lists (e.g., Maybery et al., 2002). More tellingly, grouping lists increases the tendency of participants to produce interpositions: if an item is recalled in the incorrect group, it nevertheless tends to be recalled at the correct-within group position. For example, in a 6-item list grouped into two 3-item groups, the 5th item (i.e., the second item in the second group) will, if recalled anywhere in the first three positions, tend to be recalled at the second position (that is, the second position in the first group; e.g., Henson, 1996; Lee & Estes, 1981).

The pattern of data arising from grouping is inconsistent with primacy models because those models predict that primacy will dominate in any confusions between groups (e.g., an anticipation of an item from a later group will always tend to involve the first item from that group). Although it might be argued that positional confusions occur in some other optional mechanism separate from that accounted for with a primacy gradient (Page & Henson, 2001; Page & Norris, 1998), this weakens these models as universal models of short-term order memory. Alternatively, these data have been taken to indicate a role for positional representations in short-term memory. A number of models assume that the order of items is stored by associating each item with a positional marker specifying the position of the item. By assuming that positional markers for nearby positions overlap, models incorporating positional representations are able to account for a large range of serial recall data (J. R. Anderson & Matessa, 1997; G. D. A. Brown, Neath, & Chater, 2007; G. D. A. Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Henson, 1998b; Lewandowsky & Farrell, 2008b). In particular, these models can account for the interpositions seen in grouped lists by assuming representations for both the position of items in a group, and a coarser representation of the items position in the list as a whole (item-in-list: G. D. A. Brown et al., 2000; Burgess & Hitch, 1999) or the position of the group in the list (group-in-list: J. R. Anderson & Matessa, 1997; Henson, 1998b; Lewandowsky & Farrell, 2008b). These hierarchical representations are of apparent generality, as participants may spontaneously group lists even when those lists are presented as homogeneous structures (e.g., Madigan, 1980).

Although phenomena such as grouping effects do seem to mandate some positional representations, it could be argued that incorporating such representations shifts the burden of explanation for ordering in the first place. In other words, if ordering is not a property of the items but of some external mechanism, how is that mechanism itself able to correctly order and retrieve its positional representations? The most basic answer to this question is found in models such as that proposed by Conrad (1965), who suggested that items in a sequence are stored in ordered bins in memory. Contemporary models have replaced this scheme with more detailed specifications of the generation of positional representations and their functional relationship across positions (G. D. A. Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998b). For example, an appealing mechanism for the generation of positional representations is found in the oscillator-based model of G. D. A. Brown et al. (2000), which assumes that a bank of time-varying sinusoidal oscillators generates a temporal context signal that is used to distinguish elements of a sequence and support ordered recall. In the case of grouped lists, an additional bank of oscillators is assumed to be used to code within-group position (see also Burgess & Hitch, 1999; Hitch, Burgess, Towse, & Culpin, 1996). These models, whether they specify the positional representations as varying over time (G. D. A. Brown et al., 2000) or position (Burgess & Hitch, 2006; Lewandowsky & Farrell, 2008b) can generally be classified as absolute models (Henson, 1999a) or *start-anchor* models: items are represented by their position (or time) since the beginning of a group or list. [1]

This representational scheme can be contrasted with a relative coding scheme, in which items are anchored to both the start and the end of a sequence or subsequence (cf. Braida et al., 1984). The clearest demonstration of this scheme has been in the Start-End Model (SEM) of Henson (1998b), based on the work of Houghton (1990). In SEM it is assumed that encoding the order of items consists of the storage of episodic tokens. These tokens contain information about the elements of a sequence, and also incorporate positional representations. Positional information is jointly represented by the values of a start marker and an end marker. The value of the start marker starts at some maximal value and declines across positions, thus representing proximity to the beginning of the (sub)sequence. In a complementary fashion, the end marker increases in value across positions, reaching some maximal value at the end of the sequence. When both these markers are used to represent position (see, e.g., bottom left panel of Figure 1), sequence elements are represented with respect to both the start and the end of a list. In the case of grouped lists, this relative marking scheme also applies to representation of position within groups.

Henson (1999b) presented empirical evidence for the relative marking scheme in short-term order memory. Henson ran an experiment in which participants studied 7-item lists for serial recall. The lists were presented ungrouped; with a pause between the third and fourth items (3–4 grouping); or with a pause between the fourth and fifth item (4–3 grouping). Of crucial interest was the frequency of interposition errors between positions sharing absolute within-group position, and those sharing relative within-group position. For example, in the case of the 3–4 grouping, a confusion involving the third and sixth list positions would be considered an absolute positional error, since both list positions share a within-group position of 3 (from the beginning of the group). On the other hand, a confusion of the third and seventh list items would be considered a relative error, since both positions correspond to the last item in their respective group. Henson (1999b) showed that relative errors dominated in such comparisons, even when the possibility of guessing strategies was addressed by removing reported guesses. Along with evidence that erroneous recalls of items from preceding lists (called protrusions) also tend to follow relative position, these data are consistent with the relative marking scheme assumed in models such as those of Henson (1998b) and Houghton (1990).

Although this explanation and the data offered for it are appealing, the start-end scheme comes with some theoretical baggage. As noted earlier, anchoring items to the end of a sequence presents a substantial predicament: in contrast to other domains in which such anchoring has been proposed (e.g. Braida et al., 1984; S. D. Brown et al., 2008), end-anchoring in along the temporal dimensions requires anchoring with respect to an unexperienced and unknown referent. One suggestion is that participants form some expectation about the end of the list, and use this expectation to inform the placement of the end anchor (Henson, 1999a). Although this sounds reasonable, this explanation is problematic in the case where the end of a sequence cannot be anticipated, such as when a variable list length is used in experiments (Henson, 1999a). Although the accuracy of serial recall is reduced under such conditions (e.g., Crowder, 1969; Henson, 1996, 1999b), participants are nonetheless able to recall such lists and display a normal recency effect. It is not clear in such a case how an end anchor could continue to be used without having severely detrimental effects on performance. In the specific case of Henson's (1998b) model, in which the end marker increases exponentially, under-estimating the length by only one or two positions may wreak havoc, as the end marker will continue to increase exponentially to capture the unexpected items at the end of the sequence. One possibility, explored in simulations by (Henson, 1996), is to have the end marker grow in a fixed fashion and then level off at a fixed upper limit to account for under-estimation of sequence length. Although this limits pathological behavior in the model, in fits presented by (Henson, 1996, p. 158) the model predicts that the effects of variability in list length are restricted to the final list items, whereas the data show that list length variability affects accuracy at all internal positions (Henson, 1996, p. 177).

Other mechanisms have been presented for circumventing the problem of the unknown end marker, particularly in cases where the end of a sequence or subsequence cannot be anticipated (Henson, 1996; Henson & Burgess, 1997; Houghton, 1990; see Henson, 1999a, for a discussion). For example, Henson and Burgess (1997) suggested that a number of oscillators of different frequency are initiated at the beginning of a (sub)sequence, and that the oscillator whose period gives the best match to the time scale of the sequence is then used to store the positions in that sequence. Effectively, his mechanism circumvents the problem of incorrectly estimating the length of an upcoming sequence by assuming that a large number of possible sequence lengths are considered in parallel. This mechanism is unappealing in introducing some redundancy into positional models, as each list item must be associated to all oscillators at the time of its presentation. In addition, since the oscillators are time-based, this mechanism cannot account for the insensitivity of serial recall accuracy to timing variation within lists (e.g. Lewandowsky & Brown, 2005; Lewandowsky, Brown, Wright, & Nimmo, 2006); although one explanation might be that the temporal gaps are taken to indicate the ends of groups and affect parsing of the list, the evidence suggests that such a grouping strategy cannot account for this lack of effect (Lewandowsky et al., 2006). The serial recall of grouped lists with irregular timing of items within the groups, but where overall group duration is held constant, is similarly insensitive to within-sequence timing (Ng & Maybery, 2005).

Although the results from the experiments of Henson (1999b) constitute evidence against absolute positional models, they do not necessitate the assumption of fully relative positional markers in short-term memory; not all items need be anchored to the end of the sequence. As noted by Henson (1999b), these results are equally compatible with a model in which only the last item in a group is represented with respect to the end of the group, as the critical comparisons in Henson's design all involve terminal items. A similar point was made by Page and Norris (1998) in their consideration of how grouping effects might be explained in a primacy model. Page and Norris argued that a primacy gradient was sufficient to account for the majority of phenomena in serial recall (see also Farrell & Lewandowsky, 2002), but that this gradient might be supplemented by markers representing the start, middle, and end of groups for supra-span lists (Page & Henson, 2001). In the case of groups larger than three items, all internal items might be associated with "middle" markers, with the first and last items being respectively associated with start and end markers.

In summary, the existing data suggest some form of end anchoring occurs, but does not specify how extensive this end anchoring is. Is only the last item on the list anchored to the end, or do we need to consider more complicated and theoretically burdensome explanations associated with assuming that all items are end-anchored? In the following, we use model selection to determine whether the extra complexity introduced by assuming that all items are anchored to the end of a sequence is warranted by the existing data (Henson, 1999b).

Modeling of evidence for "relative" positional representations

As a first step, we revisited the data of Henson (1999b) to determine whether the existing grouping data are sufficient to discriminate between end anchoring restricted to terminal items, and extensive end anchoring applying to all items. A set of simulations was conducted in which four models were fit to the data from Experiment 1 of Henson (1999b), which manipulated the grouping pattern of grouped lists (ungrouped, or grouped in a 3–4 or 4–3 fashion). The models, which incorporate different assumptions about the nature of positional representations in grouped lists in line with the preceding discussion, are illustrated in Figure 1.

Start-only model

The base model, termed the *start-only* model, assumed that all items in a group were anchored to the start of the group only. As illustrated in the top-left panel of Figure 1, this was accomplished by adopting the start marker from the model of Henson (1998b), with the strength of the start marker dropping across positions within the group. Although this appears similar to a primacy gradient model, items were still associated with positional representations (the single value of the start marker), such that items were cued by presenting a start marker value and matching this to the start markers of stored items; the effect of the decelerating decrease was to render items at the end of a list less distinct than those at the start. This model represents an absolute encoding scheme equivalent to the positional scheme assumed in a number of models of serial recall (Burgess & Hitch, 1999; G. D. A. Brown et al., 2000; Lewandowsky & Farrell, 2008b); the decelerating function captures the primacy within groups. Following Henson (1998b), it was assumed that a start marker was additionally used to represent the position of groups in the list as a whole. Preliminary simulations suggested that inclusion of an end marker for representing group-in-list position did not make any substantial contribution to the performance of the model; a similar observation was made by Henson (1998b), on which basis a group-in-list end marker was not included.

Restricted end model

The restricted end model builds on the start-only model by introducing an additional end marker that is associated specifically with the last item in a group. This is illustrated in the top right panel of Figure 1. All group items are associated with the continuously decreasing start marker; the end marker is only “turned on” for the last item in the group, such that only that item is anchored to the end of the group. The difference between the restricted-end model and the start-only model tells us about the additional contribution of any end anchoring. This model is in principle sufficient to handle the data of Henson (1999b), which deals only with relativity in interpositions involving terminal items.

Extensive end model

The extensive end model generalizes the restricted end model by allowing the end marker to continuously increase across the group to some maximal value in the same manner as the start marker (see bottom left of Figure 1). This scheme, which directly instantiates the positional markers assumed in the start-end model of Henson (1998b), means that all items are anchored to both the start and the end of the group. The difference between this representation scheme and the restricted end scheme addresses the central question posed here: How continuous are relative representations in short-term memory?

Primacy + start-middle-end

The final model considered here is an implementation of a variant of the restricted end marking scheme that is intended to be complementary to primacy gradient models. As suggested by Page and Norris, it is assumed that the position of items within groups is coded by generic labels for the start, middle, and end of the list. The operation of this scheme is depicted in the bottom-right panel of Figure 1. For the first position in the group, the “start” marker is activated, with the middle and end markers off. For medial positions, only the “middle” marker is activated, and for the terminal group item the “end” marker is solely activated. Additionally, the order of the items within the list as a whole is represented by a unidimensional primacy gradient of activation across all list items (not pictured in Figure 1), as assumed in primacy gradient models (Farrell & Lewandowsky, 2002; Page & Norris, 1998).

Model implementation

The four models were fit in a common framework, in a similar spirit to the lateral inhibition framework for modeling serial recall latencies employed by Farrell and Lewandowsky (2004). The purpose of this framework was to allow the comparison of the representational assumptions under controlled conditions. By keeping everything constant except for the changes of interest in the representations (with one exception for the primacy gradient model; see below), any changes in the behavior of the model can be uniquely attributed to the changes in representations. It might be objected that implementing these principles in such a framework may not reflect their "true" behavior; however, it is shown below that the behavior of the models is exactly as is would be predicted from their representational mechanisms. The simplicity of this framework necessitates a tight linkage between the representational assumptions and the predictions of the resulting model, meaning that we can with confidence reason from the model predictions back to the underlying representations.

For the start only model, restricted end, and extensive end models, items were paired with a vector of markers $\mathbf{p} = \{S, E\}$ representing the position of the item in the group, and an additional marker G representing the position of the group in the list. Following Henson (1998b), S and E were assumed to vary exponentially across the list, with the value of the markers for the jth item in a group given by

$$S(j) = \alpha^{j-1} \tag{1}$$

and

$$E(j) = \omega_0 \omega^{N_i - j}, \tag{2}$$

where N_i is the length of the group i containing the item, and α was a free parameter. An exponential relationship between distance and contextual overlap was also assumed by Farrell and Lewandowsky (2004) in their implementation of a number of a number of serial recall models in a common modelling framework, and is consistent with the fall-off in similarity with increasing distance in the models of Lewandowsky and Farrell (2008b) and G. D. A. Brown et al. (2000). The parameters ω_0 and ω were used to build up the start only, restricted end and extensive end models. For the start only model, ω_0 and ω were set to 0, meaning E was equal to 0 for all group members. For the restricted end model, ω was fixed at 0 and ω_0 was allowed to freely vary; this had the effect of setting all group items to 0, except for the last group member $E(N_i)$ which had a value of ω_0. Finally, for the extensive end model, both ω_0 and ω were left as free parameters, giving a continuous exponential across all group members. The position of the group i in the list was given by a single group marker G, also varying exponentially:

$$G(i) = \rho^{i-1} \tag{3}$$

where ρ was a free parameter. Following Henson (1998b) and Burgess and Hitch (1999), ungrouped lists were treated as lists containing a single group: group-in-list markers G were set to be identical (equal to 1), and thus did not distinguish between list items.

Recall in these three models was enacted as in positional models of serial recall (G. D. A. Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998b; Lewandowsky & Farrell, 2008b), by stepping through the positional representations in the same order as at presentation, and using the positional markers corresponding to each position to cue for an output. At each position, all list items were activated to an extent based on the overlap between their associated positional markers, and that corresponding to the position currently being cued. The overlap $o_I(c, l)$ between the positional marker of the cued position c, $\mathbf{p}(c)$, and the list item l, $\mathbf{p}(l)$, for the item-in-group markers was given by

$$o_I(c,l) = \sqrt{\mathbf{p}(c) \bullet \mathbf{p}(l)} \exp\left(-\sqrt{\sum_k (p_k(c) - p_k(l))^2}\right), \tag{4}$$

where $\bullet$ indicates the inner product between vectors (see J. A. Anderson, 1995) and k indexes the start and end components within the vector $\mathbf{p}$. Similarly, an overlap in the markers denoting group-in-list position was given by

$$o_G(c,l) = \sqrt{G(c)G(l)} \exp\left(-|G(c) - G(l)|\right), \tag{5}$$

which is a simplified version of Equation 4, given G is a scalar. Following Henson (1998b), these overlaps were used to calculate an activation value for each item l in response to cue c as

$$a(l) = o_I(c,l) o_G(c,l)(1 - r(l)), \tag{6}$$

An item was selected for recall using a probabilistic selection procedure. For the first set of simulations where group data were fit, zero-mean Gaussian noise with standard deviation σ_c (a free parameter) was added to the activations, and the item with the highest noisy activation was selected for response (see, e.g., Henson, 1998b; Lewandowsky & Farrell, 2008b; Page & Norris, 1998). This item was then suppressed by limiting its activation at following output positions (see, e.g., Farrell & Lewandowsky, 2002; Henson, 1998a; Lewandowsky, 1999; Vousden & Brown, 1998): formally, $r(l)$ in Equation 6 was set to .95 (Henson, 1998b). Response suppression is assumed in most models of serial recall, and is critical for preventing excessive perseverations in the output of the model (Henson, 1998a; Vousden & Brown, 1998).[2]

The primacy gradient model with start-middle-end marking (primacy+S-M-E) was implemented in a similar fashion to the other models. The primacy gradient across list items (for both ungrouped and grouped lists) followed an exponential function

$$S(l) = \alpha^{l-1} \tag{7}$$

where l indexes position of the item in the list. Additionally, a vector of markers $\mathbf{p} = \{S, M, E\}$ was assumed to code for position-in-group. As per Figure 1, for each position all elements were set to 0 except for that element coding the position. For the first item in a group, $\mathbf{p} = \{\tau, 0, 0\}$ (i.e., only the "start" marker was turned on); for medial items, $\mathbf{p} = \{0, \tau, 0\}$ (i.e., only the "middle" marker was turned on); and for the final item in a group, $\mathbf{p} = \{0, 0, \tau\}$. The free parameter τ ($0 \leq \tau \leq 1$) was used to effectively weight the contribution of this within-group positional representation.

Recall approximately followed the procedure in the other three models. The overlap in the within-group positional markers between the cued position c and list item l was given by

$$o_I(c, l) = \exp\left(-\sqrt{\sum_k \left(p_k(c) - p_k(l)\right)^2}\right), \tag{8}$$

k indexing the start, middle and end elements of vector $\mathbf{p}$. To implement a primacy gradient, an overlap $o_G(1, l)$ was calculated from the positional marker for item l and the positional marker for the first item:

$$o_G(c, l) = \exp\left(-|S\left(1\right) - S\left(l\right)|\right), \tag{9}$$

regardless of the actual serial position being cued.[3] This is consistent with the formulation of a primacy gradient as a gradient in the weights connecting list items to a "chunk" node, which has been implemented in a number of suggested primacy gradient models (e.g., Grossberg, 1978; Page, Cumming, Norris, Hitch, & McNeil, 2006). The activation was then calculated as per Equation 6, and the item suppressed by setting $r(l)$ to R_{init}; R_{init} was left as a free parameter given the critical role of response suppression in primacy models, leaving the primacy+S-M-E model with as many free parameters as the restricted end model.

Relation between models

The models all rely on the same fundamental equations, with the primacy+S-M-E model using a slightly different matching equation given it is driven by a primacy gradient rather than by positional representations. The start-only, restricted end and extensive end models form a series of nested models that directly address the role of various types of anchoring. The start-only model is representative of a number of models of serial recall that implement an absolute representational scheme (e.g. G. D. A. Brown et al., 2000; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008b). Although these models are already effectively rejected by the data of Henson (1999b), no fits have been reported in the literature showing that these models cannot account for those data. If these models, as represented by the start-only model, cannot account for the data (as we would expect), implementing the start-only scheme here allows us to quantify the deviation of the theory from the data; this is essential given the effects in Henson's data are numerically small. The restricted end model generalizes the start-only model by introducing an end marker restricted to the last position in a group. In turn, the extensive end model generalizes the restricted end model by allowing the end marker to extend into the entire group, and play a role in representing the position of all list markers; this model effectively implements Henson's (1998b) SEM model.

The comparison between these models directly addresses the nature of anchoring in positional representations in short-term memory. Since these are nested models, the contribution of the restricted and extensive end markers is seen by comparing them with the start-only model and the restricted end marker models, respectively, and looking at the differences between the models. If, for example, the extensive end marker introduces no benefit in accounting for the data, the extensive end marker model will behave in an identical fashion to the restricted end marker model. If the extensive end marker does introduce some important features into the model, the behavior of the model should change in a systematic fashion that better matches the patterns of positional confusions in the data. Finally, the primacy+S-M-E model addresses the plausibility of combining a primacy gradient with discrete within-group positional markers.

Modeling

The target data from the experiment of Henson (1999b) were contained in the transposition matrix for each grouping condition. For each matrix, the rows correspond to output positions, and the columns correspond to the possible input positions (i.e., position on the input sequence) that may have been recalled. The numbers in each cell of the matrix give the total number of times (aggregated across participants) that the item from each input position was recalled at each output position; the diagonal of this matrix corresponds to correct recalls.

The models were fit to the data of Henson (1999b) using maximum likelihood (ML) estimation. For a particular set of parameter values (and thus the expected probabilities produced by the model), the likelihood of the observed frequencies for a particular row (output position) is given by the multinomial likelihood function. The log-likelihood was summed across output positions and grouping conditions, and converted to a negative value to be used by the function minimization algorithm of Nelder and Mead (1965). For this first simulation, model predictions were obtained from 50,000 trial simulations. To maximize the chance of finding the global minimum for the negative log likelihood for each model, a number of starting points were used for the minimization routine. For all simulations, these points were constructed by selecting three values covering a reasonable range along each parameter, and factorially crossing the values from different parameters.

The minimized log-likelihoods were converted to Bayesian Information Criterita (BICs; Schwarz, 1978), which adjust the fit of the model to account for the bias introduced by using the same data to find parameter estimates and obtain a goodness-of-fit value. Formally, the BIC for a model is given by

$$BIC = -2 \ln \ell + 2 \ln N, \tag{10}$$

where $\ln \ell$ is the log of the maximum likelihood, and N is the number of data points (i.e., a measure of degrees of freedom) used to fit the data.[4]

Before looking at the fits in detail, the predictions of the models are shown in Figures 2 and 3 under the maximum likelihood (ML) parameter estimates. Figure 2 shows the serial position functions (SPCs) predicted by the models, with the SPCs for the data shown in the first (top left) panel. Figure 3 breaks the predictions down in finer detail by showing the fits of the models to the transposition gradients. Each panel in Figure 3 shows the difference between the transposition matrices predicted by the models, and those calculated from the data of Henson (1999b). The columns of matrices correspond to conditions, and the rows correspond to the four different models. In each panel, the difference between the observed and predicted frequency is shown for each combination of input position (i.e., position on list) and output position (where the item was actually recalled). For each combination of output position and input position, a circle indicates the size (diameter of the circle) and direction (cross: positive difference; circle: negative difference) of the difference between the model predictions and the data. For the grouped conditions, dashed lines demark the group boundaries suggested by the temporal gap.

The BIC for the start-only model was 2279.14, which is much larger (therefore indicating a worse fit) than the restricted end marker model, with a BIC of 1743.99; this BIC difference equates to extremely strong evidence in favor of the restricted end model. This shows that the addition of an end marker, albeit restricted to the final item, significantly improves the fit of the model. The second and third panels of Figure 2 show that both models do an adequate job of accounting for the data, but that the start only model tends to produce insufficient recency within groups, and in the list as a whole. Inspection of Figure 3 shows that the systematic difference between the restricted end and start only models is the underprediction by the start only model of the frequency of correct responses (on the diagonal) at the end of groups, and the overprediction of confusions between the last two items in groups (e.g., excessive confusions of items 2 and 3 in lists grouped in a 3–4 fashion).

Critically, the extensive end model, with a BIC value of 1748.06, was not superior to the restricted end model. The difference in BICs between the restricted end and extensive end models equates to a likelihood ratio of 7.77 in favor of the restricted end model. That is, when the fit of the models and their number of parameters is taken into account, the restricted end model is 7.77 times more likely to have generated the data. The evidence in favor of the restricted end model arises from a lack of difference in the models in terms of log-likelihood, and a heavier penalty applied to the extensive end model due to its extra parameter. The effective lack of difference between the models is demonstrated by the visually identical appearance of their predictions in Figures 2 and 3.

Finally, the primacy+S-M-E model had a BIC value of 2249.09. This indicates a superior in fit to the start-only model, but is far removed from the restricted and extensive end models in terms of goodness of fit. The SPCs predicted by the model in Figure 2 suggest that the primacy gradient in the model produces an excessively monotonic decrease in accuracy across serial positions. The transposition matrices in Figure 3 suggest that the difference between the primacy+S-M-E and the other models comes from its overprediction of confusions between items internal to groups of 4 items (i.e., confusions between items 5 and 6 in the 3–4 grouping condition, and between 2 and 3 in the 4–3 grouping condition).

Together, these results show that, for the data of Henson (1999b), anchoring all items to the end of a sequence (i.e., assuming a fully continuous end marker) is not necessary to enjoy the benefits of the end anchor. From the existing evidence, there is no evidence for an extensive end marker over and above a discrete end marker applying only to the terminal item.

In the following, two serial recall experiments are presented that provide further evidence on the nature and role of the end marker in serial recall. Experiment 1 examines the confusion of items internal to groups in more detail using a 4–4 grouping condition. Experiment 2 presents a within-participants replication of previous work by combining a 4–4 grouping condition (Experiment 1) with 3–4 and 4–3 grouping conditions (Henson, 1999b). The experiments are complemented by simulations quantifying the contribution of an extensive end marker in serial recall.

Experiment 1

The aim of Experiment 1 was to provide conditions favorable to the detection of an extensive end marker in serial recall of grouped lists. The major distinction between the restricted and extensive end marking schemes is that in the latter, the end marker covers all within-group positions, including positions internal to groups. Accordingly, an extensive end marker should assist in distinguishing between these internal positions. On this basis, Experiment 1 examined the positional errors arising from grouping 8-item lists in a 4–4 fashion, which allows for the examination of confusions internal to both groups, and confusions between groups involving internal items.

Method

Participants. Thirty undergraduate students from the University of Bristol participated in exchange for course credits. All participants were native or fluent English speakers. Each participant provided data for both conditions in the experiment.

Materials and apparatus. Lists were random permutations of the set of consonants *H, K, L, M, P, Q, R*, and *S*. Eighty lists were constructed for each condition (ungrouped and grouped) subject to two constraints: consecutive letters from the English alphabet could not appear in successive positions on a list (e.g., *K* and *L* could not appear in successive positions); and an item could not appear in the same serial position on consecutive lists.

The experiment was controlled by a PC that presented all stimuli (on a 17" monitor) and collected and scored all responses using the Psychophysics Toolbox for MATLAB (Brainard, 1997; Pelli, 1997).

Procedure. Participants were tested individually in a laboratory. Each trial in the experiment began with a fixation point (a cross) being presented in the centre of the screen for 1000ms. This was followed by a blank screen with duration 500ms, which was then followed by presentation of the memory list. Letters were presented one by one on the screen, each letter after the first immediately replacing the preceding item. Each letter was presented for 500ms, with a blank screen lasting for approximately 100ms being inserted between each item. Following presentation of the list there was another blank screen of 500ms, followed by presentation of the cue "RECALL" on the screen. On presentation of the recall cue, participants were to recall the letters from the list in the order they were presented by typing them on the keyboard.

In the first session of the experiment, eighty lists were presented under the conditions described so far (ungrouped lists). In the second session of the experiment, which followed the first session after a break of at least 10 minutes, participants were presented with temporally grouped lists. At the beginning of the second session, participants were informed that the eight letters would now be presented in two groups of four, and were instructed to think of the letters in two distinct groups rather than a single list. An additional pause of 500ms was inserted between the fourth and fifth list items.

The grouping instruction and additional temporal pause were introduced in the second half of the experiment for all participants. Although this leads to an order confound, this constant ordering of conditions was chosen because the grouping strategy was otherwise expected to continue once people had been presented with grouped lists, and thus contaminate the ungrouped condition (Farrell & Lewandowsky, 2004; Henson, 1999b).

Results

Although the main purpose of the data collection was to provide data for modeling, standard analyses are presented to show that the expected effects of grouping hold for these data.

Figure 4 shows serial position curves (SPCs) for accuracy (left panel) and latency (right panel). For the accuracy SPC, a 2 (grouping: ungrouped vs grouped) x 8 (serial position) repeated measures analysis of variance (ANOVA) revealed a main effect of serial position [$F(7, 203) = 107.6$, $p < .001$] and a main effect of grouping [$F(1, 29) = 25.75$, $p < .001$], with grouped lists being more accurately recalled than ungrouped lists. The interaction between grouping and serial position was also significant [$F(7, 203) = 8.34$, $p < .001$]: this is consistent with the scalloped appearance of the SPC in Figure 4 for grouped lists, suggesting primacy and recency within groups (Hitch et al., 1996). There does appear to be some non-monotonicity in the ungrouped condition suggesting some spontaneous grouping; this possibility is addressed further in the modelling.

Standard effects of grouping were also observed for the latencies (right panel of Figure 4), as measured by the time between successive keypresses (or between the recall cue and the first keypress in the case of the first item). A 2 (grouping: ungrouped vs grouped) x 8 (serial position) repeated measures ANOVA revealed a main effect of grouping [$F(1, 29) = 11.51$, $p = .002$], with grouped lists giving shorter recall times on average (mean of 1013.3 ms for ungrouped lists vs mean of 887.0 ms for grouped lists). A significant effect of serial position [$F(7, 203) = 49.04$, $p < .001$] was also observed. The interaction between grouping and serial position was also significant [$F(7, 203) = 7.58$, $p < .001$]; consistent with previous observations (e.g., Farrell, 2008; Maybery et al., 2002) a discrete peak in the latency SPC was observed for grouped lists at the group boundary (between list items 4 and 5), again indicative that standard grouping effects were obtained.

Modeling

The four models were fit to the data from Experiment 1 following the procedure used above to fit the data of Henson (1999b), with two exceptions. The first exception was a change in the modeling methodology to take advantage of the fact that individual responses were collected and retained in Experiment 1. The data were fit to individuals, rather than groups, given the possibility that a summed transposition matrix may not be representative of individuals contributing to that matrix (see, e.g., Hintzman, 1980); this was not possible for the data of Henson (1999b) as they were provided as an aggregate across participants. In obtaining fits for Henson's data reported above, it was observed that a large number of model runs were required to obtain reasonably stable probability estimates, which would preclude fitting the data of individual participants using the same procedure. To address this issue, an analytic version of the models was used, in which the noisy selection process was replaced by a generalized version of the Luce-Shepard choice function (Ashby & Maddox, 1993; Luce, 1963; Shepard, 1957):

$$p(l) = \frac{a(l)^\gamma}{\sum_i a(i)^\gamma}, \tag{11}$$

with a free parameter γ replacing the noise parameter from the earlier model. For a given set of activation values, Equation 11 gives the probability of recalling each item in that set. Henson (1998b, Footnote 2) and Page (2000) have suggested that the behavior of this choice function closely captures that of the noisy selection procedure with Gaussian noise.

A likelihood was calculated for a participant by obtaining a likelihood for each response based on the events preceding that recall on that trial. For the first output position on a trial, no items have been suppressed and the activation values are directly given by Equation 6 with all $r(l) = 0$. Following the first response, the item l that was actually recalled by the participant was then suppressed in the model by adjusting $r(l)$. The activations were then calculated for output position 2 to give a likelihood, and so on across all output positions. In the rare case of an omission or intrusion error, the response was ignored (no response suppression was applied, and no likelihood value was calculated). This procedure allows for fast and accurate maximum likelihood estimation for individual participants.

The second exception was to address the apparent spontaneous use of grouping in the ungrouped condition evident in the SPCs for Experiment 1 (Figure 4). To address this possibility, it was assumed that each response in the ungrouped condition followed from the use of the positional markers for a 4–4 grouped list with probability mix, and that the positional codes assumed for ungrouped lists were otherwise used with probability $1 - mix$. For all models the introduction of this parameter was found to significantly improve fit, and so it is presented here. The conclusions are unchanged if participants are assumed to always treat ungrouped lists as ungrouped (i.e., mix is set to 0).

Given the examination of fits for individuals in these simulations, the fitting results are presented in a slightly different format. Table 1 gives the results of the data fitting based on *model weights*. These weights are obtained by converting the corrected log-likelihoods (BICs) into likelihoods, and then scaling these so they sum to 1; formally:

$$w_i = \frac{e^{-.5BIC_i}}{\sum_j e^{-.5BIC_j}}, \tag{12}$$

where w_i is the weight for model i. When applied to the BIC values, the BIC weights can be treated as the probability that the model i is the best model for the data, and thus give a measure of the strength in favor of one model in the context of a set of models (Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004). Table 1 shows that the restricted end model (mean ML estimates: $\hat{\alpha} = .56$; $\hat{\rho} = .26$; $\hat{\omega}_0 = .29$; $\hat{\gamma} = 2.39$; $\widehat{mix} = .65$) has a distinctly greater mean weight than the other three models. Compared to the start-only model (mean ML estimates: $\hat{\alpha} = .59$; $\hat{\rho} = .26$; $\hat{\gamma} = 2.69$; $\widehat{mix} = .61$), the restricted end model predicts more recency within groups (Figure 5). Figure 6 shows that, as for the data of Henson (1999b), the major difference between the restricted end and start-only models in their transposition matrices is in the confusions involving the final and penultimate items in lists: the restricted end model predicts fewer confusions between list positions 3 and 4, and more correct responses for the fourth item.

Figures 5 and 6 show that the restricted end and extensive end models (extensive end mean *ML* estimates: $\hat{\alpha} = .55$; $\hat{\rho} = .25$; $\hat{\omega}_0 = .29$; $\hat{\omega} = .11$; $\hat{\gamma} = 2.37$; $\widehat{mix} = .63$) make almost identical predictions: the extensive end anchoring in the extensive end model appears to add little explanatory power to the model, which agrees with the much smaller number of wins and lower average BIC weight for that model in Table 1. Although the extensive end model does provide the best corrected fit in four cases, there are nearly as many wins for the start only model which assumes no end anchoring at all. Nonetheless, the average weight for the extensive end model is superior to both of the other two models. Perhaps some participants are relying substantially on extensive end anchoring? An indication that this isn't the case is seen in Figure 7, which plots the anchoring strength of internal items to the sequence ends obtained from the extensive end model. That is, by calculating the anchoring strength of the start marker to its adjacent internal item and comparing this to the anchoring strength of the end marker to its adjacent internal item, we can estimate the relative extent to which the end marker protrudes into the list, and thus determine the overall contribution of the end marker. The cluster on the right in Figure 7 shows the distribution of the strength of anchoring between the second item and the start marker (i.e., the value of the start marker at the second within-group position). These values are all fairly high (the upper bound is 1), indicating some strong anchoring between the start marker and its immediately adjacent internal item. In contrast, the anchoring values giving the strength of anchoring between the end of the group and the third item (i.e., the internal item closest to the end of the group) are clustered near 0 on the left. Although a few of these values are above 0, the largest end anchoring value is still far the below the smallest start anchoring strength seen across the participants.

Finally, the primacy+S-M-E model ($\hat{\alpha} = .85$; $\hat{\tau} = .15$; $\hat{\gamma} = 11.3$; $\widehat{R}_{init} = .63$; $\widehat{mix} = .81$) performs poorly with respect to the other three models; this model never provides a superior fit and has an average BIC weight indistinguishable from 0 (Table 1). Figure 5 shows that the SPCs predicted by the model fail to quantitatively capture the empirical SPCs. At a finer level of detail, Figure 6 shows that the model's predicted transposition matrices deviate markedly from the data, and that specifically the model overpredicts the confusion of items internal to groups (particularly items 2 and 3), as well as underpredicting performance on the fourth position.

In summary, the modeling has replicated the fits to Henson's (1999b) data: the restricted end model gives a satisfactory account of serial recall of grouped lists, and the addition of an extra parameter in the extensive end marker model is not warranted by the data. One important difference from Henson's (1999b) experiment was that the use of a 4–4 grouping structure allowed a detailed analysis of transpositions involving items internal to groups, which should have revealed any evidence for an extensive end marker having its influence across the entire group.

One possibility is that extensive end anchoring may be a general property of memory for serial order, but that this experiment was not powerful enough to test differences between the restricted end and extensive end models (despite the large differences across other model comparisons). Accordingly, an additional experiment was run to provide an additional and more generous opportunity to observe evidence for extensive end anchoring in serial recall of grouped lists.

Experiment 2

The models considered here can be distinguished not only on their predictions of accuracy, but more particularly on the patterns of transposition errors they predict. Accordingly, it seems likely that the more transposition errors are produced by participants, the more diagnostic the resulting data will be with respect to the models under consideration. In an attempt to increase the number of transpositions per participant, Experiment 2 took the strategy of running fewer people on more trials. Specifically, 8 participants were tested in four experimental sessions. Testing a smaller number of participants under multiple sessions does not reduce the power of the experiment for the purposes of modeling, since the modeling here is targeted at individual participants.

A second change attempting to increase the number of transposition errors was to introduce a parity judgement task as a distractor between list presentation and the cue to recall. Introducing a filled delay between study and test is known to reduce accuracy of serial order memory (e.g., Bjork & Healy, 1974; Farrell, 2006); given the recall set was restricted to the digits 1-9, this was expected to translate into a specific increase in the rate of transposition errors.

Finally, the ungrouped condition of Experiment 1 was replaced by the 3–4 and 4–3 grouping patterns employed by Henson (1999b). Accordingly, Experiment 2 sought to replicate the results of Henson (1999b) and Experiment 1 in a single within-participants design.

Method

Participants. Eight volunteers served as participants and were each paid 7 per session for taking part. All participants were native or fluent English speakers aged between 18-35 years. Each participant provided data for all conditions in the experiment.

Materials. Materials were as for Experiment 1, with the exception that 100 lists were constructed for each experimental session. In the 4–4 condition, lists contained eight letters. In the 3–4 and 4–3 conditions, lists were composed of seven letters.

Procedure. The procedure was identical to that for the grouped condition in Experiment 1, with a few exceptions. Participants completed four separate sessions; they took part in the 4–4 condition twice, and each of the 3–4 and 4–3 conditions once. Participants were allowed to complete no more than two sessions per day. The two 4–4 grouping sessions were presented in succession for all participants, as were the 3–4 and 4–3 conditions. The order of the 3–4 and 4–3 list sessions, and whether they preceded or followed the 4–4 lists, was counterbalanced across participants. In the 4–4 condition, a pause was added between the fourth and fifth list items to temporally group the lists; in the 3–4 condition the pause was added between the third and fourth items; and in the 4–3 condition it was between the fourth and fifth items. Finally, participants were required to make parity judgments for four digits presented individually in the centre of the screen between presentation of the list and recall. For each digit the participant was to press the left shift key in response to even digits, and the right shift key for odd digits.

Results

As for Experiment 1, standard summaries of the data are presented here only to demonstrate standard grouping effects. A full ANOVA analysis is not presented, as the design is not balanced (serial position was not fully crossed with grouping condition).

Figure 8 shows SPCs for accuracy (left panel) and latency (right panel) for Experiment 2. As in Experiment 1, it is apparent from the accuracy SPC that grouping lists affects recall, with individual primacy and recency curves present within each group in the lists. The accuracy SPCs also suggest higher accuracy for the seven item 3–4 and 4–3 lists than the eight item 4–4 lists, with means of 0.76, 0.79 and 0.66 respectively; this was confirmed by a significant one-way ANOVA [$F(2, 14) = 5.53$, $p < .05$].

The latency SPCs also illustrate the effect of grouping on recall with the pronounced peak in latencies at group boundaries (between items 4 and 5 for the 4–4 and 4–3 lists, and between items 3 and 4 for the 3–4 list).[5]

Modeling

The modeling procedure was identical to that for Experiment 1, with the exception that the mixing parameter was not required; it was clear from the data analysis that participants adopted the grouping structure suggested by the temporal pauses.

Table 1 (right) shows unanimous support for the restricted end anchoring model from the modeling results. Figure 9 shows that, as for Experiment 1, the restricted end model (mean ML estimates: $\hat{\alpha} = .58$; $\hat{\rho} = .35$; $\hat{\omega}_0 = .31$; $\hat{\gamma} = 2.31$) tends to produce more recency within groups in comparison to the start only model (mean ML estimates: $\hat{\alpha} = .62$; $\hat{\rho} = .38$; $\hat{\gamma} = 2.67$), especially within the first group. Figure 10, which shows the deviations between observed and predicted transposition matrices, also replicates Experiment 1 in revealing a greater tendency for the final and penultimate items in groups to be excessively confused in the start only model (e.g., positions 3 and 4 in 4–3 and 4–4 grouped lists).

Again, there was no evidence for a difference in fit between the restricted end model and the extensive end model (mean ML estimates: $\hat{\alpha} = .59$; $\hat{\rho} = .35$; $\hat{\omega}_0 = .359$; $\hat{\omega} = .05$; $\hat{\gamma} = 2.31$). The extensive end model had a small mean model weight, and never provided a fit superior to that of the restricted end model. As shown in Figures 9 and Figure 10, the models produced virtually identical predictions under the best-fitting parameter values, indicating that the extra parameter introduced in the extensive end model was not used to produce any changes in the performance of the model; this is also evident from the small estimate for $\hat{\omega}$.

As for the start only model, the primacy+S-M-E model ($\hat{\alpha} = .88$; $\hat{\tau} = .20$; $\hat{\gamma} = 9.51$; $\widehat{R}_{init} = .69$) failed to provide a convincing account of the data; quantitative evidence of this comes from the small model weight and low mean ranking in Table 1. As shown in Figure 9, the model overpredicts the extent of primacy within groups (compare recall at position 5 in 4–4 groups between the model's predictions and the data in Figure 8). As previously, the finer-grained breakdown in Figure 10 shows the model substantially overpredicts the frequency of confusions internal to groups.

In summary, Experiment 2 replicated the results of Experiment 1 in revealing a role for end anchoring restricted to the last item. There was no evidence of extensive end anchoring; it appears that all items bar the last are anchored only to the start of the group. Finally, the primacy+S-M-E model provided relatively poor fits to the data, suggesting that the scheme of combining a primacy gradient with nominal markers for start, middle, and end items is not a plausible solution to accounting for grouping effects in primacy gradient models.

General Discussion

The data and accompanying simulations for Experiments 1 and 2, together with the fits of the data of Henson (1999b), fail to provide support for the contribution of extensive end-anchoring to the representation of order in short-term memory. The fits to the data of Henson (1999b, involving serial recall of ungrouped, 3–4, and 4–3 grouped lists) and to the data of two new experiments (Experiment 1, involving recall of ungrouped and 4–4 grouped lists; and Experiment 2, involving serial recall of 3–4, 4–3 and 4–4 grouped lists) did not reveal any evidence for an extensive end marker over a restricted end marker that codes for only the terminal position. The extra parameter in the extensive end marker model, which weighted the anchoring of internal items to the group end, was not warranted by the minor or null improvement in fit provided by that parameter, and the predictions of the extensive end marker model were indistinguishable from those of the restricted end marker model.

Comprehensive evidence was revealed for the role of an end marker restricted to the last item in a sequence or subsequence. In all experiments, the restricted end marker model produced the overall best and most parsimonious fit; in particular, the restricted end model was found to substantially improve the fit over the model incorporating only start anchoring. Hence, although there was no evidence for fully relative representations in short-term order memory (i.e., anchoring of all items to the start and the end of the list), there is some convincing evidence for some more localized relative representations that should be incorporated into contemporary models of serial recall performance. We next discuss the implications of the modeling results for specific models of serial recall.

Implications for models of serial recall

The consistent evidence found for a restricted end marker, together with the lack of evidence for extensive end anchoring, places constraints on current models of serial order memory. The results have immediate implications for the start-end model, which was implemented here in the form of the model incorporating extensive start and end anchoring. Henson (1999b; see also Henson, 1999a) has argued for the positional representations assumed in the start-end model on the basis of his demonstration of relative confusions involving terminal items between groups of unequal size. However, as acknowledged by Henson (1999b) these data are equally compatible with the assumption of a restricted end marker confined only to terminal items. The simulations presented here have demonstrated that the cost of the extra parameter in the extensive end model is not warranted by the insubstantial improvement of fit to Henson's data, or to the data presented here.

Generally, the implication that a restricted end marker is mostly sufficient to account for serial recall performance addresses one of the core problems with the end-marking scheme in Henson's (1998b) model, which assumes that the end marker increases in an accelerating fashion across sequence positions. As discussed, having no prior knowledge about grouping structure or sequence length (e.g., Crowder, 1969) will be problematic for the model, as the end marker has the potential to explode out of bounds, or at the least interfere with relative weighting of the start and end markers (see Henson, 1996; Murdock, 2001, for a discussion). Indeed, the challenge of implementing extensive end anchoring in the face of this temporal uncertainty may explain the lack of evidence for fully relative representations despite the substantial evidence elsewhere in psychology (see Introduction).

Models that assume that positional representations are wholly absolute (G. D. A. Brown et al., 2000; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008b) have been argued unable to handle the results of Henson (1999b). The relatively poor performance of the start-only model here, in comparison with that of the restricted and extensive end models, quantifies this explanatory gap and argues for some role of an end marker in serial order memory. How might restricted end anchoring be implemented in these models? If only a restricted end marker is required, we assume that this can be "turned on" when needed (i.e., when the end of a sequence or subsequence is encountered, perhaps being triggered by a mechanism similar to that suggested by Henson & Burgess, 1997), especially in the case of temporally grouped lists, where additional time may be left for recoding of the final item. This would be straightforwardly implemented in Henson's

(1998b) SEM model in a similar fashion to that implemented here.

One model assuming absolute positional representations is the phonological loop model of Burgess and Hitch (1999) and Burgess and Hitch (2006). In their model, Burgess and Hitch assume that a window of activation slides across a set of context units, with the pattern of activation across the units at each time step giving the positional representations. Items are associated with the context states using Hebbian learning with saturation. In principle, the phonological loop model could implement a restricted end marker by reserving some units to code for the end of a list, such that these units are only activated for the last position in a sequence or subsequence (as an example, see Figure 11). In a similar fashion to the start-end model, this could be accomplished even when the sequence size is unknown for temporal grouping, as the extra time offered by the temporal gap could be used to activate the units coding the end marker, and associate these units to the item last presented. A similar mechanism could be applied to the OSCAR model of G. D. A. Brown et al. (2000), with a special "clock state" reserved to represented the end of a group or list; this clock state could then be combined (by addition or juxtaposition) with the evolving context generated by the model to represent the terminal position.

Several models of serial recall, such as those of Lewandowsky and Farrell (2008b), J. R. Anderson and Matessa (1997), and Botvinick and Plaut (2006), are agnostic about the role of the end marker in positional representations. Lewandowsky and Farrell (2008b) presented a model (see also Farrell, 2006; Lewandowsky & Farrell, 2008a), C–SOB, in which the core theoretical principles related more to the association of items with their positional representations than with the positional representations *per se*. Accordingly, they did not specify a mechanism for generating positional representations, but simply constructed the representations in line with the operation of models such as that of G. D. A. Brown et al. (2000). The suggested solutions above would all be candidates for positional representations in C–SOB. The positional representations in the ACT–R model of serial recall of J. R. Anderson and Matessa (1997) are symbolic, and are simply specified to have some similarity structure allowing for partial overlapping. In principle, the strictly ordinal representations in ACT–R could be modified to incorporate an end marker; however, the exact manner by which the overlap between this marker and other positional markers would be established is unclear. Similar comments apply to the model of Botvinick and Plaut (2006), in which items sharing within-group position were assumed to be more similar to each other than items in different within-group positions. The similarity structure of these representations was provided to the model rather than being generated by some mechanism (cf. G. D. A. Brown et al., 2000; Burgess & Hitch, 1999), so the implications of these results is not clear. Botvinick and Plaut (2006) did suggest a similar scheme to that proposed by Page and Norris (1998) in which items internal to groups share within-group position; however, the poor performance of the primacy+S-M-E model here casts doubt on the utility of this suggestion.

Indeed, these results reinforce previous conclusions that grouping effects pose a significant challenge to primacy models of serial recall (Farrell & Lewandowsky, 2002; Page & Norris, 1998). Although these models have been argued to be both sufficient (Farrell & Lewandowsky, 2002; Page & Norris, 1998) and necessary (Farrell & Lewandowsky, 2004) to account for serial recall of sub-span, ungrouped sequences, the various effects of grouping on recall errors and latencies rule these models out as comprehensive explanations for short-term order memory. The simulation results presented here also argue against the solution offered by Page and Norris (1998), who argued that the primacy gradient might be complemented by basic positional markers separate to the phonological loop. This model was generally found to perform worse (or equal worst) than the other models (particularly the two models incorporating an end marker) in the model comparisons presented above. It may be that there is some other implementation of the model that would perform better than the version implemented here; however, the systematic excess of confusions internal to groups containing four items points to a likely fundamental flaw in the assumption of a "middle" marker common to all internal positions.

One remedy for primacy models would be to assume that both fully-developed positional markers (as in the models of G. D. A. Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998b) and a primacy gradient are used to represent and recall the order of elements in a sequence. On the one hand, this would undermine the simplicity of these models, and would introduce a problem of credit assignment for these models: if they are successful in accounting for some serial recall phenomena, to what extent is this success due to the primacy gradient rather than the positional markers? On the other hand, a primacy gradient may be required in addition to positional markers to comprehensively explain serial recall phenomena. Most positional models incorporate a primacy gradient, whether it be in decreasing learning rates across the list (G. D. A. Brown et al., 2000; Farrell, 2006; Lewandowsky & Farrell, 2008b) or allowing for saturation of weights by earlier list items (Burgess & Hitch, 1999). Futhermore, Farrell and Lewandowsky (2004) have argued that a primacy gradient (complemented by response suppression) is needed to explain the pattern of latencies for positional errors, whereby anticipation errors (recalling later items at earlier output positions) tend to be slower than postponements (recall earlier items at later output positions). Arguably, the best account offered for the data of Farrell and Lewandowsky (2004) comes from the model of Lewandowsky and Farrell (2008b). Lewandowsky and Farrell's model is a dynamic distributed model in which items are associated with positional representations, and in which the strength of these associations is controlled by an energy-gating mechanism which effectively reduces the learning rate across list positions (i.e., a primacy gradient).

Other evidence for relative representations

Additional evidence for relative representations in serial recall comes from an additional experiment conducted by Henson (1999b) in which he examined protrusion errors, whereby items from previously presented lists are erroneously recalled on the current list. As for confusions between groups, Henson (1999b) found that confusions between lists involving the terminal item were made on the basis of relative position; for example, if a 5-item list was followed by a 7-item list, protrusions involving the 5th (last) item of the preceding list would tend to appear at the 7th (last) rather than 5th position in the next list. Although a natural explanation for these protrusion errors is that serial order representations are re-used between lists, recent evidence suggest that order information is only re-used if lists are repeated verbatim (Cumming, Page, & Norris, 2003; Hitch, Fastame, & Flude, 2005; Page et al., 2006). As for the case of grouping effects (Page & Norris, 1998), this has led some to downplay the contribution of positional representations to serial recall (Cumming et al., 2003). Regardless, if it is assumed that positional markers are replicated between lists, the results of Henson (1999b) suggest a role for some relative marking in serial recall. Given the simulations above did not reveal a role for extensive end marking in Henson's grouping experiment, there is some cause to question whether the results for the protrusion experiment—also examining confusions of terminal items—dictate fully relative end markers. Because the focus here has been on single-trial effects (these being the target of the majority of serial recall models to date), we leave this as an open empirical question.

A final note on spontaneous grouping

In closing, one reason for examining the role and nature of the end marker in serial recall is the importance of positional markers in explaining grouping effects generally. The natural role for grouped representations in short-term memory is reinforced by suggestions that participants spontaneously group lists, even in the absence of any objective group cues (Farrell, 2008; Henson, 1996; Madigan, 1980). This spontaneous grouping manifests itself as a "kink" in the serial position function similar to that observed in explicitly grouped lists, but less extreme (see, e.g., Figure 5), and has been observed across a number of studies (Madigan, 1980). In one sense, this spontaneous grouping might provide a realistic mechanism for extensive end anchoring, by assuming that participants a priori determine a grouping pattern for ungrouped lists and can therefore set up an appropriate end anchor. However, the fits from these three experiments have shown no evidence for an extensive end marker in the recall of grouped lists, again disputing the use of extensive end anchoring.

One benefit of the modeling of Experiment 1 is that an estimate of spontaneous grouping was provided by the parameter mix, which quantifies the weighting of grouped positional representations in accounting for ungrouped lists. The mean value of mix across participants for the best-fitting model was .65, indicating widespread spontaneous application of grouped representations to ungrouped lists. Figure 12 shows a breakdown of this parameter across participants in the form of a histogram. This histogram confirms the prevalence of spontaneous grouping in Experiment 1,[6] and also highlights the variability between participants, with some showing very little evidence of spontaneous grouping, and other participants showing behavior indicative of a comprehensive use of grouped positional markers for ungrouped lists. What is not clear, and is a target for future research, is whether this variability reflects individual differences in strategic use of grouping (see, e.g., Towse, Hitch, & Skeates, 1999) or more fundamental differences that may be tapped by spontaneous grouping. In any case, it is clear that grouped representations are a fundamental aspect of short-term memory.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723.

Anderson, J. A. (1995). *An introduction to neural networks.* Cambridge, MA: MIT Press.

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review*, *104*, 728-748.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology*, *37*, 372-400.

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

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

Braida, L. D., Lim, J. S., Berliner, J. E., Durlach, N. I., Rabinowitz, W. M., & Purks, S. R. (1984). Intensity perception: XIII. Perceptual anchor model of context-coding. *Journal of the Acoustical Society of America*, *76*, 722–731.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Bressan, P. (2006). The place of white in a world of grays: A double-anchoring theory of lightness perception. *Psychological Review*, *113*, 526–553.

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.

Brown, S. D., Marley, A. A. J., Donkin, C., & Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychological Review*, *115*, 396–425.

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

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

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd ed.). New York: Springer-Verlag.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding aic and bic in model selection. *Sociological Methods & Research*, *33*, 261-304.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascade model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204-256.

Conrad, R. (1965). Order error in immediate recall of sequences. *Journal of Verbal Learning & Verbal Behavior*, *4*, 161–169.

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

Cumming, N., Page, M., & Norris, D. (2003). Testing a positional model of the hebb effect. *Memory*, *11*, 43-63.

Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 923–941.

Duncan, M., & Lewandowsky, S. (2005). The time course of response suppression: No evidence for a gradual release from inhibition. *Memory*, *13*, 236-246.

Farrell, S. (2006). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language*, *55*, 587–600.

Farrell, S. (2008). Multiple roles for time in short-term memory: Evidence from serial recall of order and timing. *Journal of Experimental Psychology: Learning, Memory,* *and Cognition*, *34*, 128–145.

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

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*, 115-135.

Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*, *9*, 5-22.

Gravetter, F., & Lockhead, G. R. (1973). Criterial range as a frame of reference for stimulus judgment. *Psychological Review*, *80*, 203–216.

Grossberg, S. (1978). A theory of human memory: Self organization and performance of sensory-motor codes, maps and plans. In R. Roden & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5). New York: Academic Press.

Henson, R. N. A. (1996). *Short-term memory for serial order.* Doctoral dissertation, University of Cambridge.

Henson, R. N. A. (1998a). Item repetition in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1162-1181.

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

Henson, R. N. A. (1999a). Coding position in short-term memory. *International Journal of Psychology*, *34*, 403-409.

Henson, R. N. A. (1999b). Positional information in short-term memory: Relative or absolute? *Memory & Cognition*, *27*, 915-927.

Henson, R. N. A., & Burgess, N. (1997). Representations of serial order. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *Proceedings of fourth neural computation and psychology workshop* (p. 283-300). London:: Springer.

Hintzman, D. L. (1980). Simpson's paradox and the analysis of memory retrieval. *Psychological Review*, *87*, 398-410.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology*, *49A*, 116-139.

Hitch, G. J., Fastame, M. C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory*, *13*, 247-258.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation.* London: Academic Press.

Hsee, C. K., Hastie, R., & Chen, J. (2008). Hedonomics: Bridging decision research with happiness research. *Perspectives on Psychological Science*, *3*, 224–243.

Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: A response-time analysis. *Psychological Research*, *60*, 121–133.

Lawless, H. T., Horne, J., & Spiers, W. (2000). Contrast and range effects for category, magnitude and labeled magnitude scales in judgements of sweetness intensity. *Chemical Senses*, *25*, 85–92.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 149–169.

Lewandowsky, S. (1999). Redintegration and response suppression in serial recall: A dynamic network model. *International Journal of Psychology*, *34*, 434-446.

Lewandowsky, S., & Brown, G. D. A. (2005). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory*, *13*, 283-292.

Lewandowsky, S., Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language*, *54*, 20-38.

Lewandowsky, S., & Farrell, S. (2008a). Phonological similarity in serial recall: Constraints on theories of memory. *Journal of Memory and Language*, *58*, 429–448.

Lewandowsky, S., & Farrell, S. (2008b). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 49, pp. 1–48). London, UK: Elsevier.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103–189). New York: Wiley.

Madigan, S. (1980). The serial position curve in immediate serial recall. *Bulletin of the Psychonomic Society*, *15*, 335-338.

Maybery, M. T., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory and Language*, *47*, 360-385.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Murdock, B. (2001). Analysis of the serial position curve. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder.* Washington, DC.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, *7*, 308-313.

Ng, H. L. H., & Maybery, M. T. (2005). Grouping in short-term memory: Do oscillators code the positions of items? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 175-81.

Page, M. P. A. (2000). Connectionist modelling in psychology: A localist manifesto.

Behavioral and Brain Sciences, *23*, 443–467.

Page, M. P. A., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 716–733.

Page, M. P. A., & Henson, R. N. A. (2001). Computational models of short-term memory: Modelling serial recall of verbal material. In J. Andrade (Ed.), *Working memory: A work in progress* (p. 177-198). London: Routledge.

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

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437-442.

Reeves, C., Schmauder, A. R., & Morris, R. K. (2000). Stress grouping improves performance on an immediate serial list recall task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1638-1654.

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology*, *21*, 137-147.

Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology*, *21*, 148-155.

Sadalla, E. K., Burroughs, W. J., & Staplin, L. J. (1980). Reference points in spatial cognition. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 516–528.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461-464.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*, 325-345.

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgement. *Psychological Review*, *112*, 881–911.

Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, *53*, 1-26.

Towse, J. N., Hitch, G. J., & Skeates, S. (1999). Developmental sensitivity to temporal grouping effects in short-term memory. *International Journal of Behavioral Development*, *23*, 391-411.

Vousden, J. I., & Brown, G. D. A. (1998). To repeat or not to repeat: The time course of response suppression in sequential behaviour. In D. G. J.A. Bullinaria & G. Houghton. (Eds.), *Proceedings of the fourth neural computation and psychology workshop: Connectionist representations* (pp. 301–315). London: Springer-Verlag.

Wagenmakers, E.-J., & Farrell, S. (2004). Aic model selection using akaike weights. *Psychonomic Bulletin & Review*, *11*, 192-196.

Wickelgren, W. A. (1964). Size of rehearsal group and short-term memory. *Journal of Experimental Psychology*, *68*, 413-419.

Author Note

This research was supported by ESRC grant RES-062-23-0272. We are indebted to Rik Henson for providing the data from his grouping experiments, and his permission to report the modeling of those data here. Portions of this work were presented at the January 2008 meeting of the Experimental Psychology Society, London, UK, and the Seventh Annual Summer Interdisciplinary Conference, Madonna Di Campiglio, Italy, July 2008. Correspondence should be addressed to Simon Farrell, Department of Psychology, University of Bristol, 12a Priory Road, Clifton, Bristol BS8 1TU, UK; e-mail: Simon.Farrell@bristol.ac.uk.

Notes

[1]Henson (1999a) also discusses the possibility of having all items represented by their absolute position with respect to the end of a group or list. Since no contemporary models make this assumption it is not discussed here.

[2]Response suppression was assumed not to decay across output positions (cf. Henson, 1998b). There is evidence against the notion that response suppression drops off over time (Duncan & Lewandowsky, 2005), and restricting response suppression to be fixed across positions limited the number of free parameters.

[3]The inner product multipliers in Equations 4 and 5 were omitted for the primacy+S-M-E model, as these were found to excessively weight earlier list items, and thus produce too steep a decline in performance across serial positions.

[4]One choice that needed to be made here was between using the AIC (Akaike's Information Criterion; Akaike, 1974) and the BIC as a corrected log-likelihood. There is no "right" choice between these metrics, with each being appropriate for different purposes. BIC was used here on the basis that some of the models were nested, and on the basis that the profile of log-likelihoods was not tapered, but showed some big differences and some small differences between models; see Burnham and Anderson (2004) for further details.

[5]The data were also inspected for consistency with the results of Henson (1999b). Specifically, the proportions of transpositions involving the critical positions in 3–4 and 4–3 grouped lists was calculated from weighted log-odds following Henson (1999b). The proportion of total number of errors that were absolute and relative transpositions was 0.11 and 0.12 for 3–4 grouped lists, and 0.10 and 0.15 for 4–3 grouped lists, respectively. These cannot be analyzed as per presentation in Henson (1999b) as an ungrouped condition was not included here. Nonetheless, in both cases the proportions of relative transpositions exceeded those of absolute transpositions.

[6]This can be contrasted with the conclusions of Murdock (2001), who found little evidence of chunking (i.e., grouping) when fitting a simple model which assumed binomial distributions of chunk sizes and probability of recall given chunk size. Farrell and Lewandowsky (2002) argued that this questioned the prevalence of spontaneous grouping for short lists. However, Murdock (2001) did not benchmark his model against explicitly grouped lists, and it is unclear whether his model would correctly produce the serial position functions usually taken as evidence for spontaneous grouping (e.g., Madigan, 1980). Nonetheless, it seems reasonable to suggest that a grouping structure is less likely to be enforced for shorter lists.

Table 1

BIC statistics of fits of the four models to Experiment 1 (left) and Experiment 2 (right). The columns in each half give the mean BIC weight (M); the standard deviation in the BIC weights (SD); the mean rank of for each model when the models are ranked on the basis of BIC for each individual; and the number of times each model was the "best" model (ie., had the largest BIC model weight).

	Experiment 1				Experiment 2			
Model	*M*	*SD*	*M* Rank	Wins	*M*	*SD*	*M* Rank	Wins
Start only	0.10	0.26	2.77	3	0.00	0.00	3.00	0
Restricted end	**0.73**	0.39	1.30	**23**	**0.95**	0.09	1.00	**8**
Extensive end	0.17	0.30	1.97	4	0.05	0.09	2.00	0
Primacy + S-M-E	0.00	0.00	3.30	0	0.00	0.00	3.13	0

Figure Captions

Figure 1. Four possible schemes for representing position in groups. Top left: Start only marker, in which all group items are represented with respect to the start of the group; Top right: Restricted end marker, in which all items are represented with respect to the start of the group, and the last group item is additionally associated with a specific end marker; Bottom left: Extended end marker, in which all items are represented with respect to the start and the end of the group; Bottom right: Start-middle-end, in which the first group item is associated with a start marker, the final group item is associated with an end marker, and internal items are associated with a common middle marker. In all panels, S indicates a start marker, M a middle marker, and E an end marker.

Figure 2. Serial position functions for the fits of four models to the data of Henson (1999b).

Figure 3. Transposition matrix differences between the maximum likelihood predictions and the data of Henson (1999b). Circles containing a cross indicate a positive difference; open circles indicate a negative difference. The size of the circle indicates the absolute size of the discrepancy. See text for further description.

Figure 4. Serial position functions for Experiment 1. Left panel: Mean proportion correct by serial position, for ungrouped and grouped conditions; Right panel: Mean recall latencies by serial position.

Figure 5. Serial position functions for the fits of four models to the data from Experiment 1.

Figure 6. Transposition matrix differences between the maximum likelihood predictions and the data from Experiment 1 for all participants.

Figure 7. Distributions across participants of the anchoring strength of the second item with respect to the start of the sequence (dark grey), and the third (penultimate) item with respect to the end of the sequence (light grey).

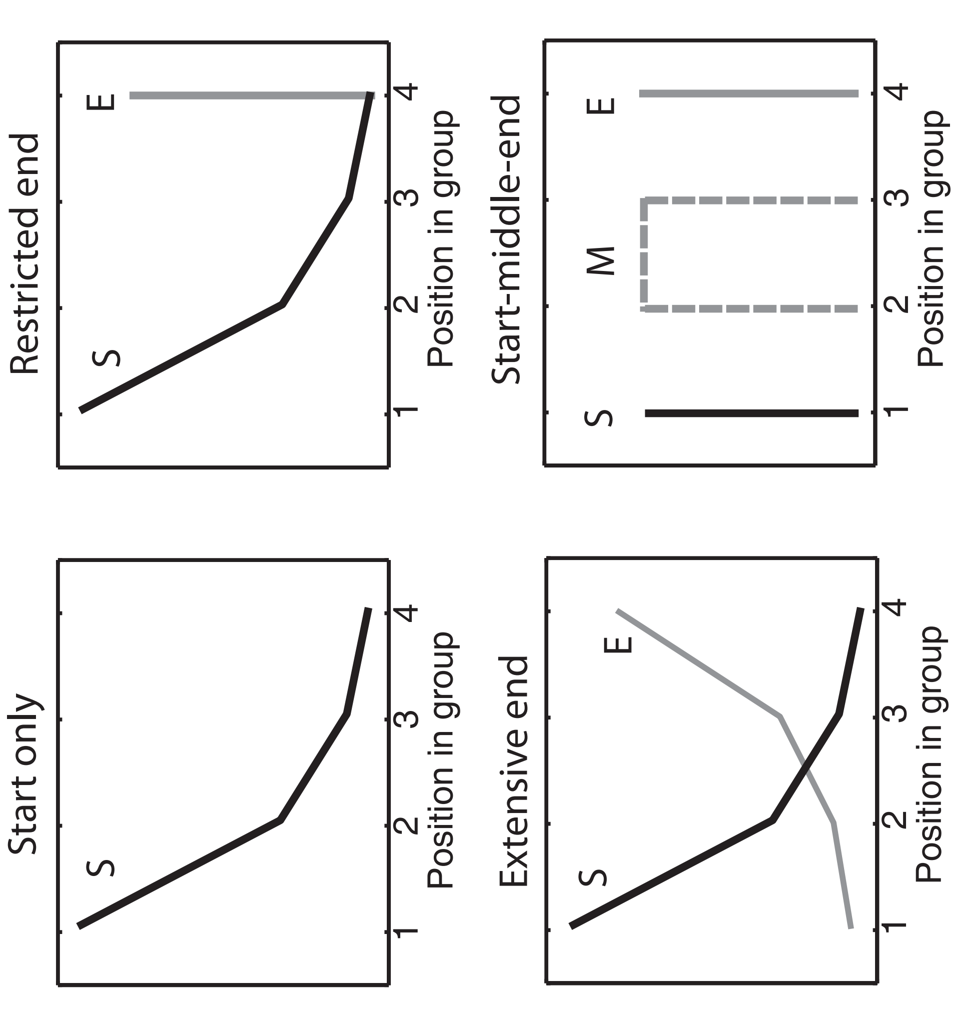
Figure 8. Serial position functions for Experiment 2. Left panel: Mean proportion correct by serial position, for ungrouped and grouped conditions; Right panel: Mean recall latencies by serial position.

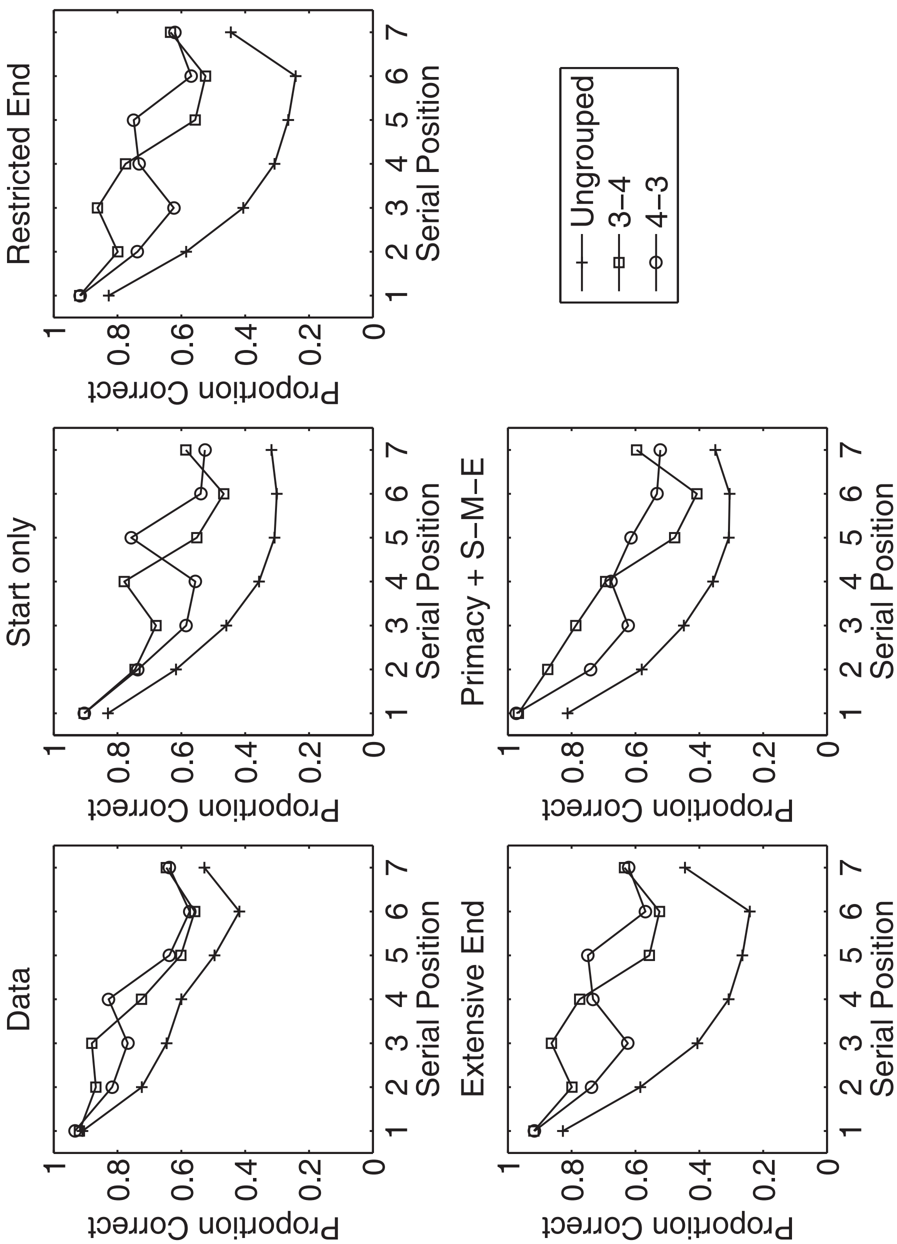
Figure 9. Serial position functions for the fits of four models to the data from Experiment 2.

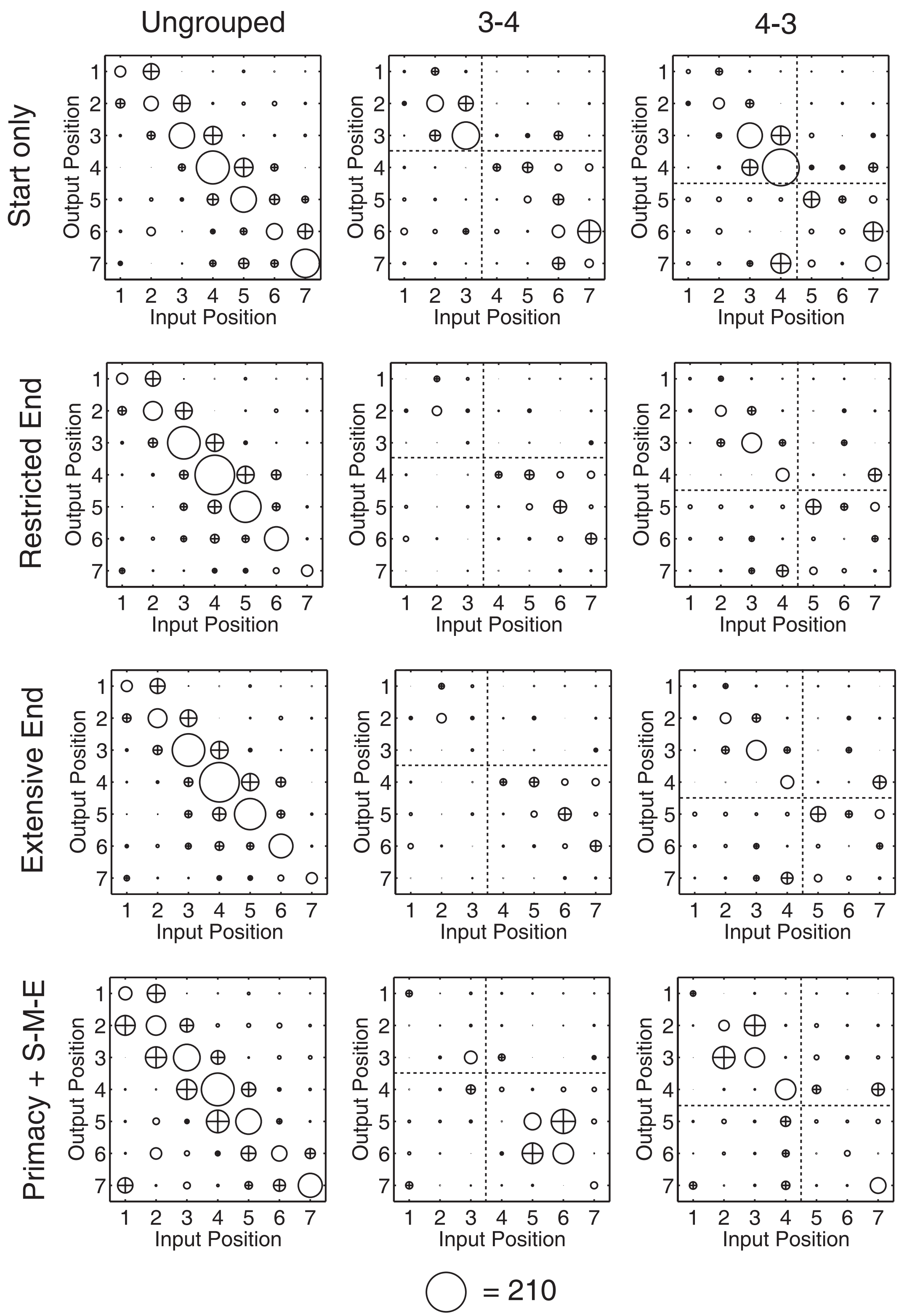
Figure 10. Transposition matrix differences between the maximum likelihood predictions and the data of Experiment 2.

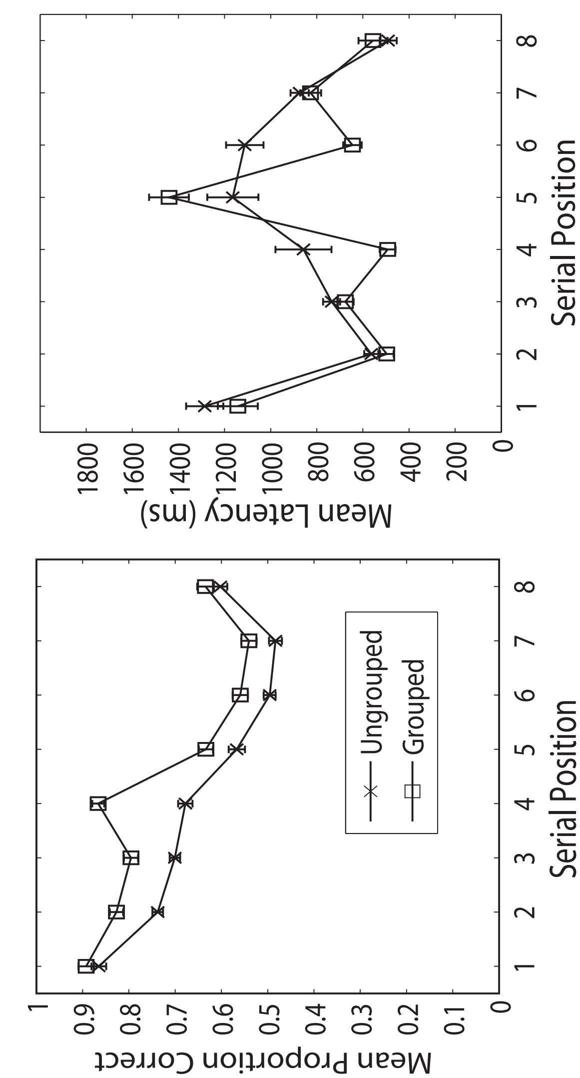
Figure 11. A restricted end-marking scheme for the Burgess & Hitch (1999, 2006) model. A sliding window of activations is used to code the position of items with a group or sequence, with a dedicated set of "end" nodes being activated only for the terminal item.

Figure 12. Histogram of *mix*, the parameter quantiying the extent of spontaneous grouping in ungrouped lists, for Experiment 1.

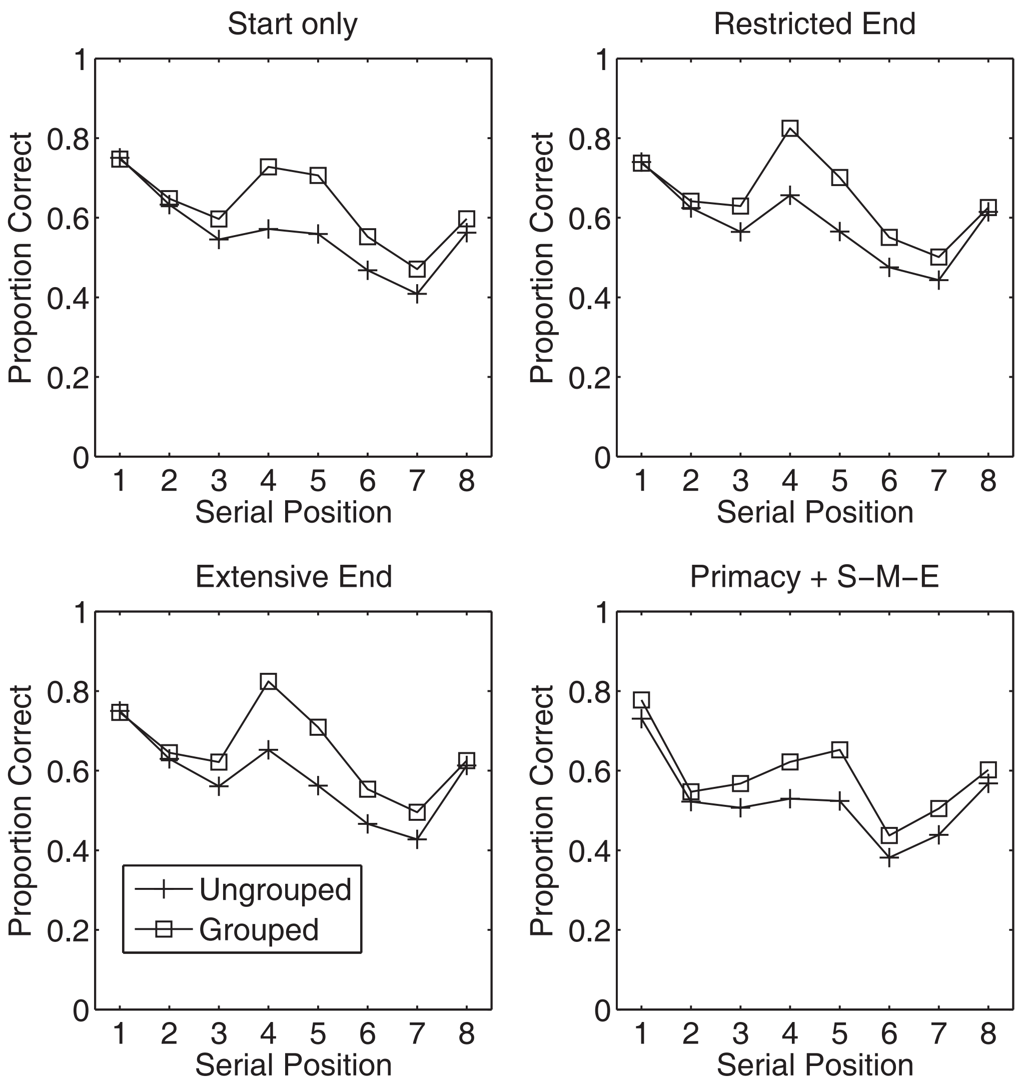

Start only
Restricted end
Extensive end
Start-middle-end
S
E
M
Position in group
Strength of marker
1
2
3
4

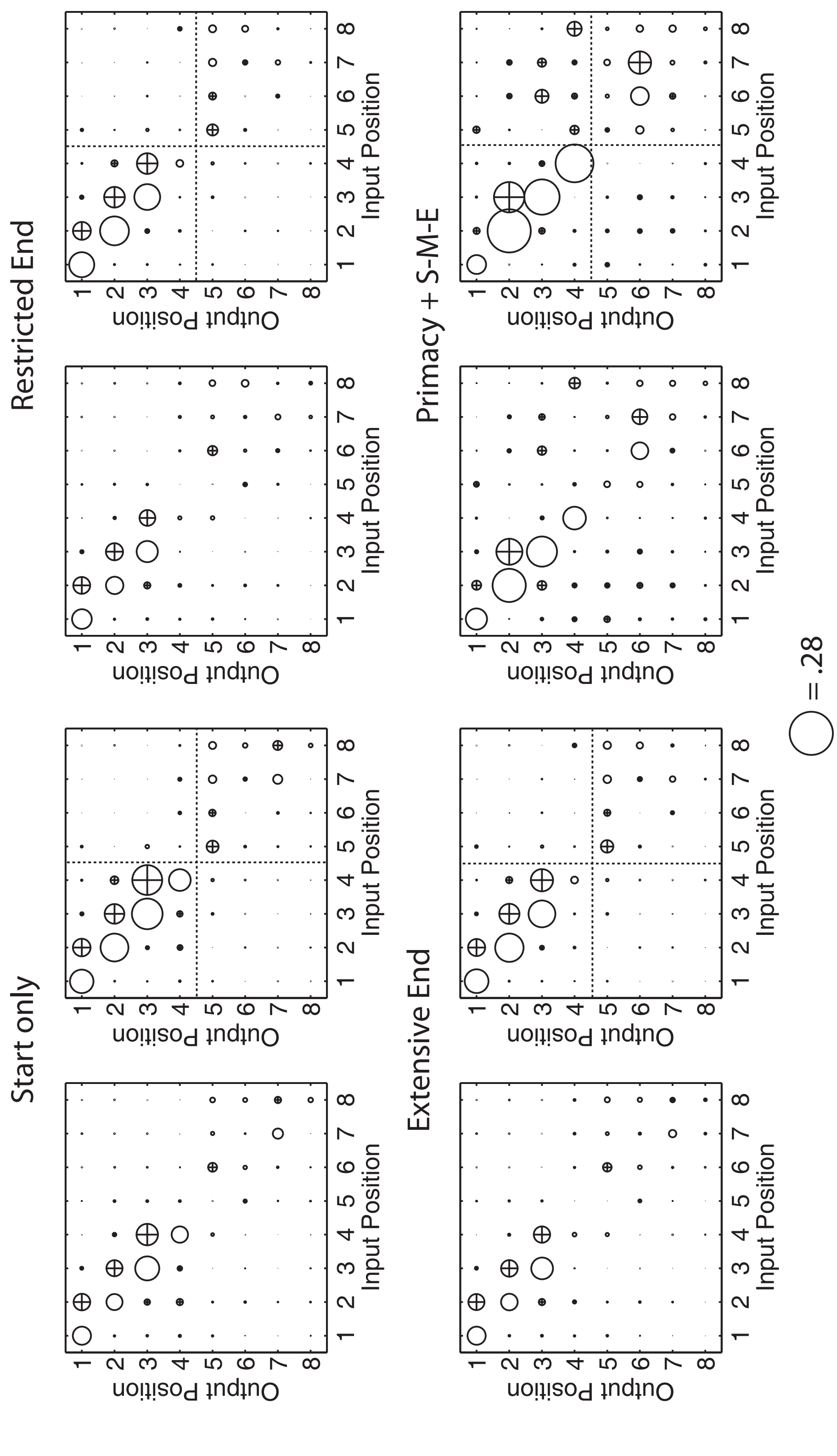

Data
Start only
Restricted End
Extensive End
Primacy + S–M–E
Proportion Correct
Serial Position
Ungrouped
3–4
4–3

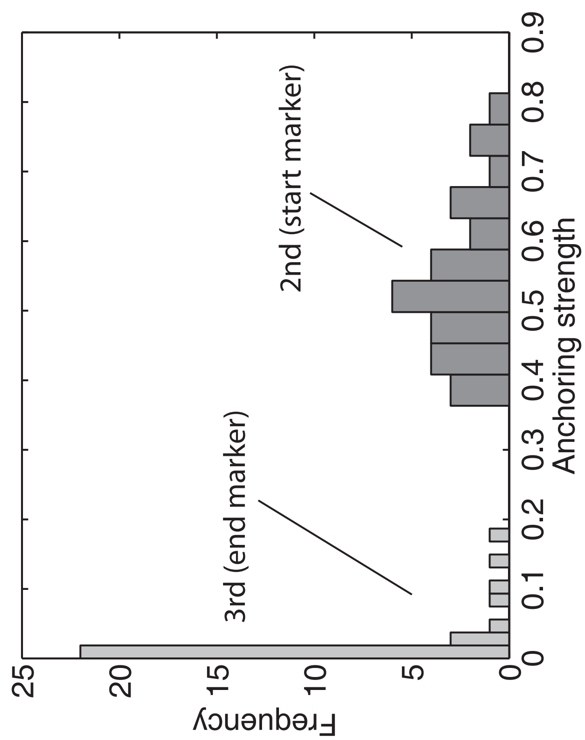

Ungrouped
3-4
4-3
Start only
Restricted End
Extensive End
Primacy + S-M-E
Output Position
Input Position
1 2 3 4 5 6 7
= 210

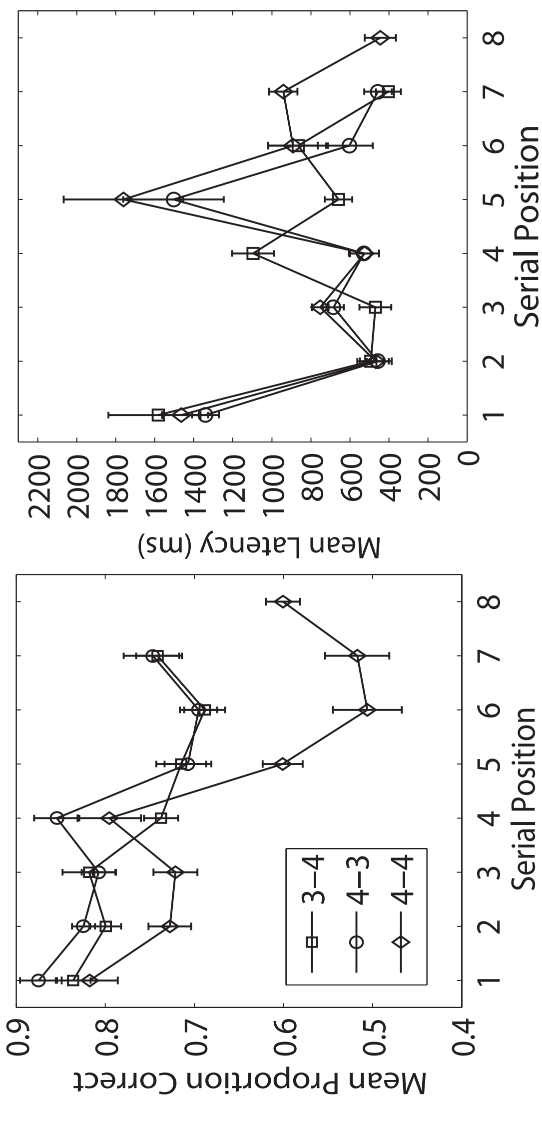


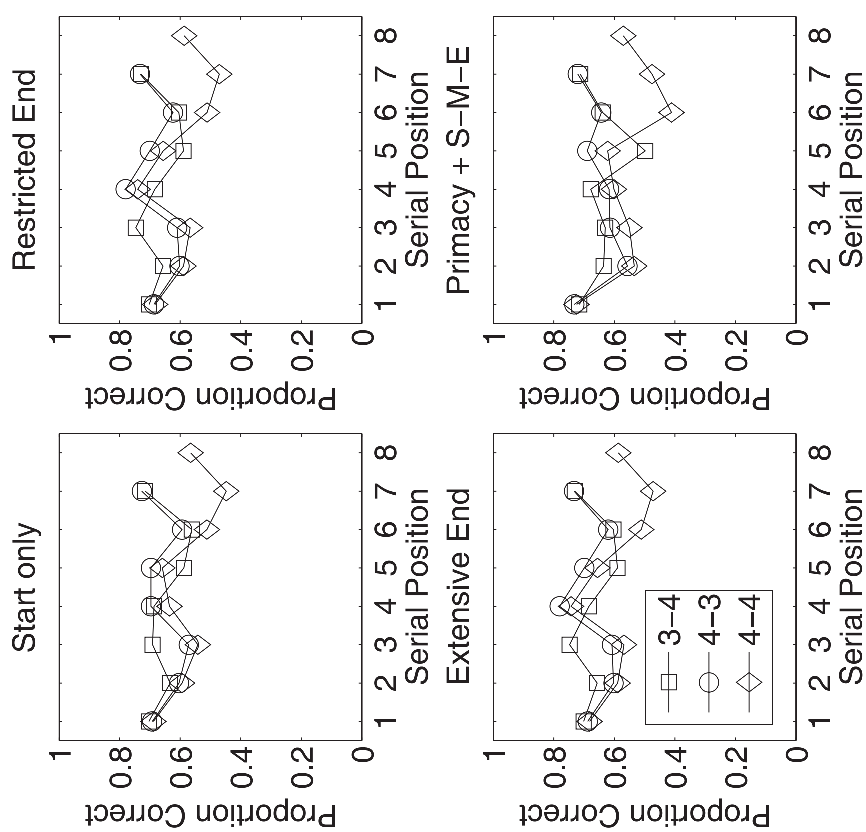
Mean Proportion Correct
Serial Position
Ungrouped
Grouped
Mean Latency (ms)
Serial Position

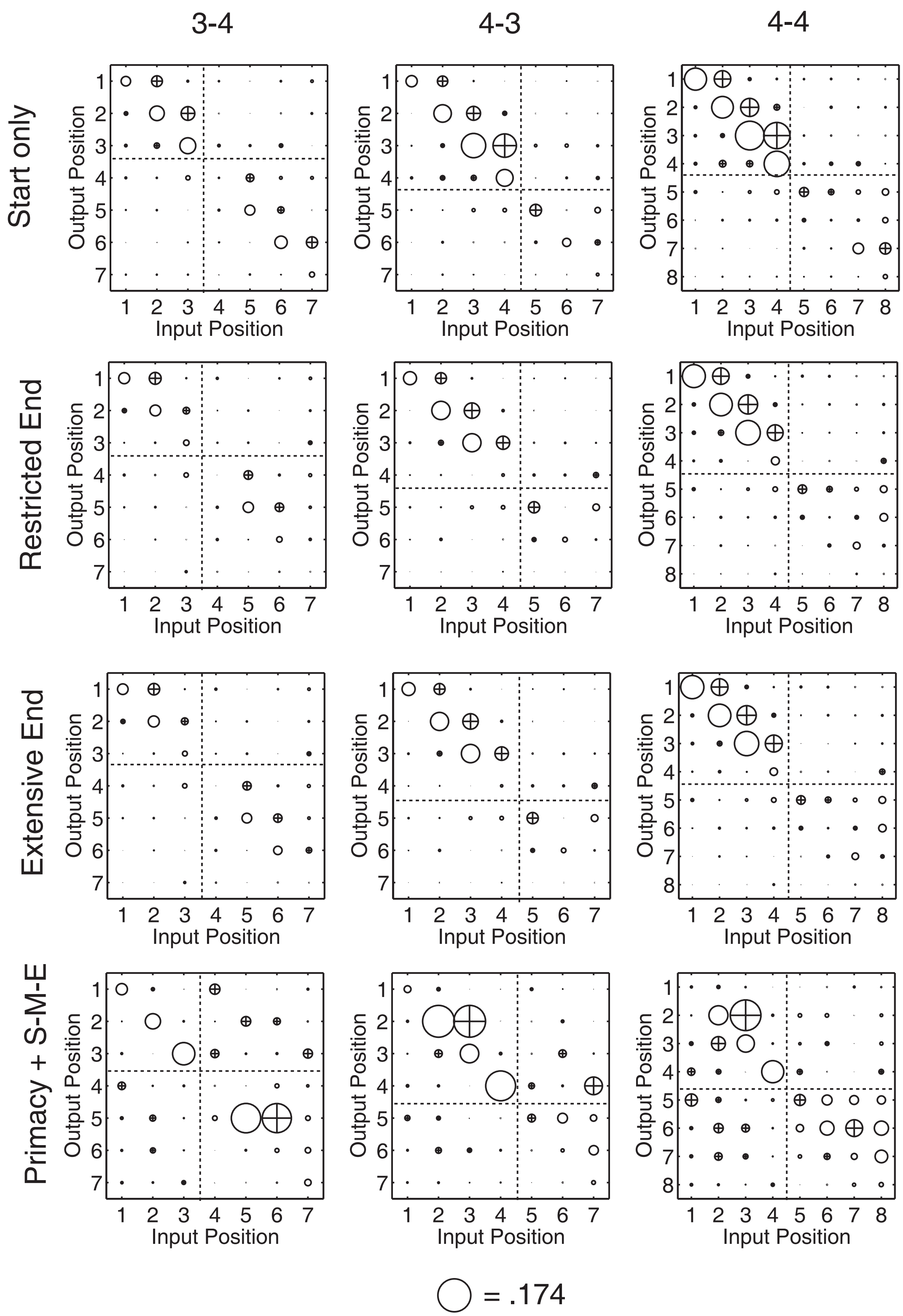

Start only
Restricted End
Extensive End
Primacy + S−M−E
Proportion Correct
Serial Position
0
0.2
0.4
0.6
0.8
1
1
2
3
4
5
6
7
8
Ungrouped
Grouped

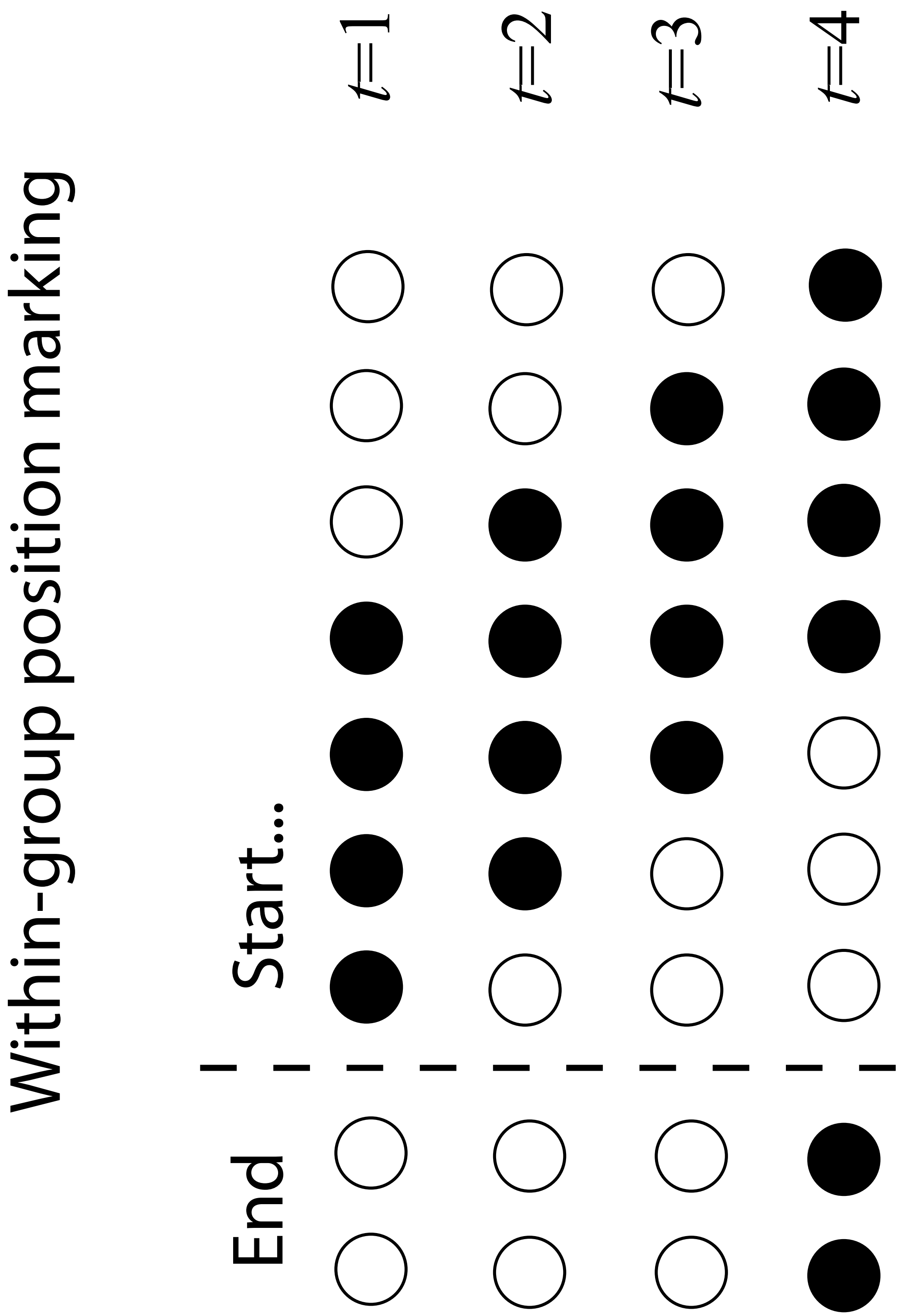

Start only
Restricted End
Extensive End
Primacy + S-M-E
Output Position
Input Position
= .28


Frequency
Anchoring strength
3rd (end marker)
2nd (start marker)
0
5
10
15
20
25
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Mean Proportion Correct
Serial Position
3–4
4–3
4–4
Mean Latency (ms)
Serial Position


Start only
Restricted End
Extensive End
Primacy + S–M–E
Proportion Correct
Serial Position
3–4
4–3
4–4


3-4
4-3
4-4
Start only
Restricted End
Extensive End
Primacy + S-M-E
Output Position
Input Position
= .174


Within-group position marking
t=1
t=2
t=3
t=4
Start...
End

End anchoring in STM, Figure 12