Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism

Tom Hartley[a*], Mark J. Hurlstone[a,b], Graham J. Hitch[a]

[a]Department of Psychology, University of York, Heslington, York, UK, YO10 5DD

[b]Research Section (Psychology), University of Western Australia (M304), 35 Stirling Highway, CRAWLEY WA 6009, Australia

Corresponding Author :

Tom Hartley
Department of Psychology
University of York
Heslington
York
UK
YO10 5DD
tom.hartley@york.ac.uk
+44 1904 322903

Key words: short-term memory; timing; serial order; speech perception; speech production

Abstract

Immediate memory for spoken sequences depends on their rhythm – different levels of accuracy and patterns of error are seen according to the way in which items are spaced in time. Current models address these phenomena only partially or not at all. We investigate the idea that temporal grouping effects are an emergent property of a general serial ordering mechanism based on a population of oscillators locally-sensitive to amplitude modulations on different temporal scales. Two experiments show that the effects of temporal grouping are independent of the predictability of the grouping pattern, consistent with the model's stimulus-driven mechanism and inconsistent with alternative accounts. The second experiment reports detailed and systematic differences in the recall of irregularly grouped sequences that are broadly consistent with predictions of the new model. We suggest that the bottom-up multi-scale population oscillator (or BUMP) mechanism is a useful starting point for a general account of serial order in language processing more widely.

Effects of rhythm on memory for spoken sequences: a model and tests of its stimulus-driven mechanism

1 Introduction

Language is inherently serial, and the representation and control of serial order are fundamental considerations for any model addressing linguistic processes or the interactions of memory and language. Over half a century ago, K.S. Lashley (1951) recognized that the capacity for serial behavior had to be reconciled with evidence of parallel processing in the brain, and suggested that rhythmic patterns of neural activity could play a mediating role. Since then theoretical approaches to serial order within the language domain have tended to diverge, with a proliferation of models for speech production (e.g., Dell, Burger & Svec, 1997; MacKay, 1970; Vousden et al., 2000), verbal short-term memory (e.g., Botvinick & Plaut, 2006; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Henson, 1998; Lewandowsky & Farrell, 2008; Page & Norris, 1998), and speech perception (e.g., Grossberg, 2003). In each area models have incorporated many of Lashley's (1951) insights, but typically not that of the central and potentially unifying importance of rhythm.

In the present paper we focus on effects of rhythm and timing in auditory-verbal short-term memory. Recall of rhythmically grouped sequences is typically much better than ungrouped sequences, and the improvement is associated with characteristic changes in patterns of order errors (Ryan, 1969a, 1969b). These phenomena place constraints on theories of serial order in short-term memory and these in turn have implications for developing an understanding of the broader problem of serial order in language more generally. We replicate and extend the work of Ryan (1969a) showing grouping effects for irregular and unpredictable patterns of temporal grouping which, we argue, are inconsistent with explanations of grouping in terms of strategic processes such as rehearsal (Broadbent, 1975; Chi, 1976; Lewandowsky & Brown, 2005; Parmentier & Maybery, 2008; Wickelgren, 1964; 1967), and beyond the scope of current computational models of serial order in short-term memory. Using insights from such models, we propose a new mechanism in which serial order is encoded by a population of oscillators driven bottom-up by auditory-verbal input and sensitive to local variation in its temporal structure. Through simulations we demonstrate that many subtle and detailed features of the empirical data on short-term memory for grouped sequences can be understood as emergent properties of this general mechanism. We conclude by discussing the potential of the bottom-up multi-scale population oscillator (or BUMP) mechanism as a starting point for a more general theory of serial order in language processing, potentially linking speech perception, speech production and verbal short-term memory.

1.1 Verbal short-term memory: structure, function and mechanisms

Although it might initially appear an esoteric skill, the capacity to retain ordered spoken material over a brief interval is fundamental for many aspects of language processing. Thus, many authors view verbal short-term memory as a property of the language processing system (e.g., Allen & Hulme, 2006; Martin & Saffran, 1997; Monsell, 1987), and the capacity for immediate verbal recall has been shown to play a key role in the acquisition of vocabulary and language development (Baddeley, Gathercole, & Papagno, 1998). The mechanisms underpinning verbal short-term memory have principally been studied in tasks involving the immediate serial recall of sequences of items such as digits, letters or words. There is substantial evidence from performance in such tasks that the underlying system is speech-based. Thus, recall is disrupted when items sound alike (Conrad, 1964) or take longer to say (Baddeley, Thomson, & Buchanan,

1975) or when the memory task is accompanied by irrelevant spoken output (Murray, 1967). These effects of phonological similarity, word length and articulatory suppression fall into a systematic pattern that has been widely interpreted as reflecting the operation of a store containing transient phonological memory traces that can be refreshed by subvocal rehearsal (Baddeley & Hitch, 1994; see also Baddeley, 1986; 2007). Despite its critics (e.g. Jones, Hughes & Macken, 2007), this account of verbal short-term memory as a “phonological loop” has been highly influential. However, in its original form it offered no explanation of memory for serial order and did not address key phenomena such as the shape of serial position curves, the distribution of order errors, and effects of temporally grouping items during sequence presentation. The following sections describe these phenomena and the computational models that have attempted to explain them. (for a fuller treatment of data and models, see Hurlstone, Hitch, & Baddeley, 2014).

1.2 Serial order effects in verbal short-term memory

The most basic phenomenon is the serial position curve plotting the accuracy of recalling items as a function of their position in a sequence. When items are presented at regular intervals the curve is characteristically bow-shaped reflecting a tendency for items at the beginning and end of the sequence to be better recalled than items in the middle. The most common error in recalling a sequence of familiar items is a transposition, where an item is recalled at an incorrect position. These order errors obey the locality constraint (Henson 1996), in that adjacent transpositions are most common, and the frequency of transpositions decreases as displacement increases. However, different results are seen when sequences are temporally grouped during presentation by inserting extra pauses, typically after every third item. Recall is substantially more accurate due largely to a reduction in order errors, and the serial position curve is multiply-bowed, with a bow for each group (Frankish, 1985; Hitch, Burgess, Towse, & Culpin, 1996; Ryan, 1969a, 1969b). The reduction in order errors is largely due to a reduction in the frequency of adjacent transpositions, especially across group boundaries (Henson, 1999; Maybery, Parmentier, & Jones, 2002). However, the frequency of one type of between-group transposition is increased by grouping: these are interposition errors (Henson, 1996), between-group transpositions that preserve their within-group positions (Farrell & Lelievre, 2009; Henson, 1999; Ng & Maybery, 2002; 2005; Ryan, 1969b). It is easy to appreciate that these effects of temporal grouping provide a rich set of constraints on possible mechanisms for serial ordering, and this has indeed proved to be the case.

1.2 Models of serial order

In recent years there has been considerable progress in developing computational models of verbal short-term memory that offer mechanistic accounts of serial order and the associated empirical phenomena (see e.g., Botvinick & Plaut, 2006; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Farrell, 2012; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Farrell, 2008; Lewandowsky & Murdock, 1989; Page & Norris, 1998). We provide a brief overview of the main features of these models below, and we show how temporal grouping effects help discriminate amongst them.

Models of memory for serial order can be regarded as falling into two broad categories depending on whether order is coded via associations between different items or, alternatively, between items and a separate context signal. According to chaining models (Lewandowsky & Murdock, 1989; Murdock, 1995), order is represented through associations between successive items. Serial recall is accomplished by retrieving the first item, which then cues the second item, which cues the third, and so on and so forth. Chaining is intuitively appealing but has difficulty accounting for the locality constraint on order errors and paired transpositions, the tendency for items close together in a list to swap places. Since the chaining mechanism only activates forthcoming items, it cannot readily explain how an earlier item can take the place of a later one as an error. Chaining is also incompatible with sawtooth error patterns found in serial recall of lists comprising alternating phonemically similar and dissimilar items (Henson, 1996; Henson, Norris, Page, & Baddeley, 1996)[1].

The other, currently dominant, class of models makes the assumption that each item is associated with the current state of an internal context signal that changes gradually during sequence presentation. To recall the sequence, the context signal is replayed and items are reactivated in parallel according to the similarity between the current state of the signal and the state to which each item was associated. This process establishes an activation gradient across the items, such that the correct item is most active, followed by its near neighbours (which are associated with more similar states of the signal than more distant neighbours). Context signal models typically involve a competitive queuing (CQ) process that translates the simultaneous pattern of activations into a sequential output, whereby the most active item is selected, output, and then immediately suppressed. This leaves the remaining items to compete for selection when the replay of the context signal moves on to the next serial position, and so on. Errors are modelled by assuming activation levels contain a certain amount of noise. Context signal models can comfortably account for the locality constraint on transposition errors as neighbouring items compete more strongly than distant items. Indeed local transpositions are a characteristic property: if the wrong item is output at one serial position it is likely to come from a neighbouring position, while the correct item will remain in competition at the next serial position. Context models also provide a natural account for the bowed shape of the serial position curve, as items at the start and end of a sequence have fewer competing neighbours than items in the middle.

In the simplest models, the context signal is a monotonic marker that is most active at the beginning of a list and becomes progressively weaker thereafter, as in the Primacy Model of Page and Norris (1998). However, such a simple marker cannot readily deal with situations in which the same item is repeated at different positions. This limitation has been addressed by assuming a more complex, multidimensional context signal. For example, Houghton (1990) and later Henson (1998) assumed a two-element context signal in which an additional end-of-list marker becomes increasingly active as the start-of-list marker's activity dies away. Such a signal represents approximate position relative to the start and end of the sequence. This added complexity deals with the problem of lists containing repeated items but is insufficient to account for temporal grouping effects. The characteristic combination of multiply bowed serial position curves, reduced adjacent transposition errors and increased interposition errors has been taken to imply that grouped sequences are organised hierarchically (see e.g., McNicol & Heathcote, 1986). To accommodate grouping effects, models typically assume a more complex multidimensional context signal. For example, Henson's (1998) Start-End model assumes that temporally grouped lists recruit an additional two-element context signal that codes position relative to the start and end of each group. The combination of these two signals has the dimensions of relative position-in-list and relative position-in-group, and has the necessary similarity structure to reproduce the multiply bowed serial position curves and error patterns typically observed in recall of temporally grouped lists. The higher overall level of recall associated with grouping arises straightforwardly from the increased discriminability of states of the context signal when it contains two dimensions rather than one.

Other models differ chiefly in assuming a context signal that reflects either absolute position from the start of the sequence, as in C-SOB (Lewandowsky & Farrell, 2008) and the model originally proposed by Burgess and Hitch (1992), or a temporal correlate of absolute position, as in OSCAR (Brown, Preece & Hulme, 2000) and later versions of the Burgess and Hitch model (Burgess & Hitch, 1999, 2006). OSCAR is particularly interesting with regard to the present investigation. This model assumes that the context signal reflects the combined outputs of a set of free-running, endogenous oscillators pre-set to run at different rates. To illustrate this, Brown et al. (2000) use an analogy of the rotating hands on a clock face. Thus, for a regularly grouped list, the seconds, minutes and hour hands might correspond to the rates of presenting individual items, groups and whole lists. States of the clock face would be more discriminable from one another for a grouped than an ungrouped list, due to the presence of the extra hand, but by the same token would go through a similar series during the presentation of each group. Once again, therefore, the hierarchical structure of the context signal accounts for the different patterns of recall observed for ungrouped and regularly grouped lists described earlier.

Regardless of which specific approach is adopted, all context signal models account for temporal grouping effects by assuming that the signal changes in such a way as to combine information about position-in-list with position-in-group. However, none of the models specifies the mechanism through which groups are detected and the within-group component of the context signal is reset. For example, Burgess and Hitch (1999) assumed that the context signal reflects internal oscillators entrained to the rhythms of items and groups, but the processes of entrainment were not implemented and the context signal was altered “by hand” to simulate the recall of grouped and ungrouped lists. This omission is also a characteristic of the SEM (Henson,

1998) and C-SOB models (Lewandowsky & Farrell, 2008). Similarly, in OSCAR (Brown et al., 2000) it was simply assumed that grouped presentation recruits an additional repeating component of the context signal and the processes whereby this takes place were left unspecified. Indeed, Brown et al. (2000) were careful to note that they had not modelled the recruitment of oscillators of appropriate frequency, and that it may not be straightforward to do this. Thus, while all the models succeed in reproducing the effects of regular, predictable grouping patterns on serial recall, they beg the question of how it is that the context signal changes in response to a rhythmic input (or indeed what processes determine the context signal for an ungrouped input).

The absence of detailed implementation is an important limitation as a genuinely useful model will generate insights into the representation of serial order for all kinds of verbal sequence, including natural sequences, where the rhythms are typically irregular and vary in their predictability. Present models clearly lack the information needed to do this. An interesting partial exception is the multi-oscillator model developed by Henson and Burgess (1997) in an attempt to account for the dependence of interposition errors on relative rather than absolute (i.e. temporal) within-group position in sequences containing groups of unequal size (Ng & Maybery, 2005). These effects challenge models in which the context signal has a strictly hierarchical structure based on absolute temporal position. To address them, Henson and Burgess (1997) made use of two important ideas. One is that oscillators are arranged in pairs with the same frequency but different phases 90° apart. The other is that oscillator pairs compete in parallel to represent the input. Henson and Burgess (1997) used these assumptions to develop a model that was able to reproduce not only the classic effects of regular temporal grouping but also the subtle effects for sequences containing groups of unequal size reported by Ng and Maybery (2005).

However, Henson and Burgess (1997) noted limitations of their model in that they had to assume prior rhythmic parsing of the input and did not specify how the phases of different oscillator pairs were synchronised. This lack of implementation detail means it is not possible to predict how the model will behave when presented with unpredictable, irregularly grouped sequences that cannot be parsed in advance.

The model we present here develops insights in the approaches taken by Henson and Burgess (1997) and Brown et al. (2000), but addresses their limitations and those of the other context signal models we have been discussing. We do this by meeting the challenge of specifying a genuinely "bottom-up" processing mechanism whereby the context signal is computed from the information in the physical stimulus, as it unfolds in real time. Rather than assuming a set of free-running oscillators (Brown et al., 2000, Henson & Burgess, 1997), which as we have seen raises the unanswered question of how the relevant oscillators for any sequence are recruited, we take a fresh approach to the problem by turning it around and assuming instead a set of frequency-sensitive detectors that respond to local rhythms in the stimulus input. We show that a context signal based on the outputs of an array of such detectors has the desired properties and generates predictions for human data in the recall of irregular or unpredictably grouped sequences. We go on to describe experiments reporting human data and evaluate the degree to which our bottom-up serial ordering mechanism can account for them.

To summarize: the most successful models explaining temporal grouping effects assume that serial order is retained via associations between items and a context signal that changes progressively during the presentation of a sequence. Although debate has understandably tended to focus on differences between models regarding the precise nature of the context signal, a wider concern is that they have concentrated on explaining only a limited set of data on the effects of temporal grouping and currently none of the models is capable of explaining the recall of sequences grouped irregularly or unpredictably. This is a significant problem because, in the general case, and especially in more ecologically realistic circumstances, we cannot anticipate the rate or temporal structure of the words and phonemes we encounter and whose serial order we must represent. Models capable of representing sequences with arbitrary timing will necessarily be more general, and may provide new insights into language processing outside the laboratory where the rate and structure of speech we perceive and produce is neither regular nor predictable. In the next section we outline a model that addresses this limitation by implementing a context signal that is sensitive to local aspects of sequence timing.

1.4 Current Model

The current model's basic architecture is common to most models of this type, as outlined above. Thus, during presentation, each item is associated with the current state of a temporal context signal. At recall, successive states of the temporal context signal are replayed, and the learned associations reactivate the items. At each step items compete for selection on the basis of their levels of activation and are then suppressed, so that the selected item is not available for retrieval at the next serial position.

The model makes use of established signal processing mechanisms to analyze the temporal structure of an auditory input, and assumes a set of neurally-plausible bottom-up detectors sensitive to local changes in the amplitude envelope of incoming speech. The responses of these hypothetical neurons combine to form the temporal context signal. The property of bottom-up, local sensitivity is crucial for the model's sensitivity to unpredictable, irregularly grouped input sequences as well as predictable regular temporal grouping patterns. For brevity, we call this the BUMP (standing for *Bottom-Up Multi-scale Population*) mechanism, and we use the BUMP model, to refer to its application to the simulation of auditory-verbal serial memory.

In the BUMP model, the context signal reflects the activity of a population of cells, each with an intrinsic tuning which causes its activity to oscillate at a specific rate and phase in response to local changes in the speech envelope (Figure 1a). This pattern of activity is determined by the neuron's characteristic impulse response, which can be thought of as an ideal signal against which the incoming amplitude modulations are compared. More precisely, the activity of each model neuron is determined by convolving its impulse response with the input signal. If driven by more or less regular fluctuations in the amplitude of the speech signal at a rate close to its intrinsic tuning, the amplitude of the neuron's response will be strong and, importantly, its phase will track that of the incoming stimuli. If driven at a markedly different frequency, the amplitude of the response will be much lower. For example, a given neuron might be tuned to respond maximally to regular modulations at 2Hz in phase with peaks in the speech envelope (Figure 1b, left). Critically, each neuron's activity will continue to track the phase of amplitude modulations as their rate varies over a substantial range around its ideal tuning, albeit with somewhat reduced amplitude (Figure 1b, right) but beyond this range the amplitude of the neuron's response falls away. Each neuron thus responds only to local amplitude modulations at a specific temporal scale and neurons with similar tunings tend to respond in a coherent way to amplitude modulations on the relevant scale.

Figure 1 here

In the current implementation we use phase-offset pairs of model neurons tuned to the same frequency, each pair corresponding to a complex quadrature filter. For each frequency one member of the pair has a maximum response aligned with peaks in the speech envelope (0°

phase) and the other has a maximum response offset by 90° (Figure 2a). The response of such a pair to a sequence of items is determined by the amplitude modulations associated with its presentation, which we approximate as a series of triangular pulses. Figure 2b illustrates the responses of a single pair of neurons with tunings close to the group repetition rate during presentation of three groups of three items, as in Ryan (1969b). For easy visualisation, the offset responses of each pair have been combined to illustrate the overall pattern of amplitude and phase responses during sequence presentation with phase represented by hue and amplitude represented by brightness (Figure 2c). Figure 2d shows that the phase of the output tracks local changes in the speech envelope corresponding to groups of items, that is, the phase indicated by the joint response of the pair goes through one complete cycle for each group of three items in the input sequence, while the overall amplitude of the response remains more or less constant throughout. An important feature of the model is that pairs of neurons tuned to different frequencies will show different patterns of output, as described below.

Figure 2 here

In BUMP, the context signal used to encode a verbal sequence is the output of a population of oscillator pairs whose tunings span the range of temporal scales over which amplitude modulations are encountered in typical tasks (Figure 3). For example in a short, irregularly-timed list of discrete items, there may be slow modulations on a temporal scale corresponding to the length of the entire utterance (say 5s), more rapid fluctuations corresponding to groups or clusters of items (perhaps around 1-2s), and still faster modulations corresponding to presentation of the individual items themselves (0.75s). In the current simulations we use 15 pairs of oscillators in the range 0.1-1.3 Hz so as to cover such fluctuations.

The outputs of oscillator pairs tuned to local amplitude fluctuations on different time-scales will vary during presentation of a list, determined by its temporal structure at different time-scales. In the simplest example of an evenly-timed, ungrouped list, oscillators with tunings close to the item presentation rate will respond strongly and in phase with the items. These oscillators will go through one cycle per list item. Oscillators with tunings close to the list presentation rate respond to the larger scale amplitude fluctuation associated with presentation of the entire list and are insensitive to the relatively rapid changes associated with individual items. These slower oscillators' output will go through approximately half a cycle during presentation of the list. Oscillators with intermediate tunings respond only weakly and their responses are largely restricted to the beginning and end of the sequence. However, for a regularly grouped list, we have already seen that oscillators with tunings close to the group presentation rate are also recruited, and go through one cycle per group (see Figure 2d). A critical advantage of the bottom-up mechanism in the BUMP model, and in particular the sensitivity of the context signal to *local* changes in the input, is that it can also deal with irregularly grouped stimuli. These features are illustrated in Figure 3 which compares the context signal generated for a) an ungrouped list of nine items, b) a list with an evenly spaced 3-3-3 grouping pattern and c) an uneven 2-6-1 pattern.

Figure 3 here

In the BUMP model, grouping tends to be advantageous because it activates a broader population of oscillators, resulting in a more distinctive representation of serial position that reduces competition between neighbouring items at retrieval, as seen qualitatively in the comparison between Figure 3a and 3b. An important property of the model is that some temporal grouping patterns are more favourable than others. Regular grouping patterns (3-3-3) powerfully activate oscillators tuned to the grouping rate, which enhances overall recall, albeit at the expense of interposition errors (as seen in the similarity of the outputs of filters tuned to the group presentation rate in Fig 3b). Irregular grouping patterns similarly favour intergroup transpositions although the correspondence between different positions is less clear-cut (see Figure 3c). The inconsistency of group durations means that oscillators in this range are less strongly activated than would be the case for regularly grouped lists. Context signals associated with different list positions are, on average, less distinctive, though some serial positions (especially those in very short groups or at the beginning or end of longer groups) may be protected relative to those in ungrouped or regularly-grouped lists.

Overall, the BUMP model predicts that certain grouping patterns will be more favourable for serial recall than others. We report two experiments testing the model. The first addresses its basic assumption that the benefit of auditory-verbal grouping depends on bottom-up processing of information about the temporal structure of the speech envelope as distinct from top-down encoding strategies based on an expected structure (Experiment 1). We then compare the specific predictions of the model with empirical data on the recall of irregularly grouped lists for a wide range of different grouping patterns (Experiment 2), and with data from Ryan (1969a).

2 Empirical tests of BUMP's bottom-up mechanism

As we have seen, a major assumption of the BUMP model is that the state changes of the putative timing signal responsible for ordering in serial recall are driven by bottom-up, stimulus-based properties, rather than top-down expectations regarding the list structure. One way of testing this assumption is by exploring whether the effects of temporal grouping are sensitive to the availability of foreknowledge of the grouping structure on an upcoming trial.

We have already noted the widely held assumption that grouping effects in immediate serial recall reflect the implementation of optional strategies such as rehearsal (Broadbent, 1975; Chi, 1976; Lewandowsky & Brown, 2005; Parmentier & Maybery, 2008; Wickelgren, 1964; 1967). This is plausible when one considers that almost all documented studies of temporal grouping have employed either the same pattern of grouping throughout the entire experiment (Farrell & Lewandowsky, 2004; Frankish, 1985, 1989; Henson, 1996; Hitch et al., 1996; Maybery et al., 2002; Parmentier & Maybery, 2008; Ryan, 1969b) or a small set of different patterns presented in blocked fashion (Farrell & Lelievre, 2009; Henson, 1999). In such cases participants may be able to capitalize on their foreknowledge of the locations of the pauses and use them to engage in extra rehearsal of the preceding group, thereby strengthening within-group positional codes.

An important but neglected exception to the above generalization is a study by Ryan (1969a; Experiment 2) in which participants received auditorily presented lists of nine digits. Although lists were always presented in three temporal groups, these varied in size to give 28 different grouping patterns that varied unpredictably from trial to trial. Even though participants could not anticipate the pattern of grouping on a forthcoming trial, there were systematic differences in recall with serial position curves exhibiting scalloped profiles specific to and consistent with the pattern of grouping in question. These results suggest that the genesis of grouping effects—in the auditory modality at least—is attributable to a bottom-up ordering mechanism. Nevertheless, Ryan's results do not preclude the possibility that top-down processes based on foreknowledge of the list structure might modulate the effects of grouping. Thus, it is possible that when the list structure is predictable—as would be the case in a version of Ryan's experiment in which the different grouping patterns were presented in separate blocks of trials—

the effect of a particular pattern of grouping on serial recall might be stronger. However, since Ryan did not incorporate such a condition in her study, the answer to this question is presently unclear.

In the following experiments we sought to obtain further evidence for the contention that temporal grouping effects in the auditory modality are attributable to a bottom-up ordering mechanism. We accomplished this by manipulating the predictability of the grouping structure of temporally grouped lists.

2.1 Experiment 1

We compared the effects of temporally grouping spoken lists of nine digits into threes when advance knowledge of the grouping structure either was or was not available. In one condition (predictable 3-3-3), participants were always presented with lists that were temporally grouped into threes. In a second condition (unpredictable 3-3-3), a separate set of participants received lists in which there was wide trial-to-trial variability in the grouping pattern, rendering it virtually impossible to predict the grouping structure on a forthcoming trial. Crucially, lists were presented in groups of threes on only a randomly distributed 20% of trials. In a further baseline condition (ungrouped), another set of participants received lists that were devoid of any grouping cues. The three conditions were compared by contrasting the recall of lists grouped in threes in the unpredictable 3-3-3 condition with recall on corresponding experimental trials in the predictable 3-3-3 and ungrouped conditions.

If advance knowledge of the grouping structure is necessary for the manifestation of grouping effects then these effects should be absent in the unpredictable 3-3-3 condition. If, on the other hand, advance knowledge simply modulates rather than underpins the effects of grouping, then these effects should be present in both grouping conditions, but stronger in the predictable condition. If, however, grouping effects are present in both conditions in equal magnitude then this would support the view that such effects have a bottom-up locus.

One potential problem of interpretation is that participants might have a default strategy of subjectively grouping lists into threes. Subjective grouping is a common strategy with visually presented ungrouped lists (see e.g., Farrell & Lelievre, 2009; Henson, 1996; Madigan, 1980) and might be expected to occur with spoken lists too. The problem for interpretation is that if participants were to subjectively group *all* lists into threes then performance in the predictable and unpredictable 3-3-3 conditions would be very similar, rendering it difficult to determine whether this is due to the bottom-up nature of grouping or the ubiquity of a subjective 3-3-3 grouping strategy. We attempted to avoid such ambiguity by including conditions in which any tendency to group subjectively using rehearsal would be disrupted by requiring articulatory suppression during list presentation (see e.g., Lewandowsky & Brown, 2005). If the effects of grouping in the predictable and unpredictable 3-3-3 conditions are comparable, then the conclusion that grouping is due to bottom-up processes would be bolstered by the survival of this pattern under articulatory suppression.

2.1.1 Method

Participants. Thirty members of the campus community at the University of York took part in the experiment in exchange for course credit or payment of £4.

Stimuli & Apparatus. The stimuli were spoken lists consisting of the digits 1-9. Lists were generated quasi-randomly according to the following constraints: (1) no ascending or descending runs of more than two digits, (2) no adjacent digits in immediately neighbouring positions, (3) no repetitions of the same digit, and (4) no digits in adjacent lists sharing the same within-list position. Digits were recorded in a monotone male voice with a 16-bit resolution at a sampling rate of 22, 050 kHz using Sony Sound Forge 8.0 software. They were normalized to 0dB, compressed to a duration of 400ms, and saved as sound files on a Dell Dimension 3100 PC. Stimulus presentation was controlled by the same computer equipped with a 17 inch monitor using software developed in-house. The digits were presented via Senheisser HD 265 Linear headphones at a sound level of approximately 65dB.

Design. The experiment employed a 3 × 2 mixed factorial design in which list-type (ungrouped vs. predictable 3-3-3 vs. unpredictable 3-3-3) was a between-participant factor, and secondary task (no suppression vs. suppression) was a within-participant factor. Counterbalancing was employed on the latter factor. Ten participants were randomly allocated to each of the three between-participant conditions.

Blocks of 50 lists were constructed as follows. In the ungrouped condition, items were presented at a uniform rate, whilst in the predictable and unpredictable 3-3-3 conditions two extended temporal pauses served to demarcate lists into three groups. In the predictable 3-3-3 condition, the extended pauses were always located after the third and sixth items so as to create lists that were perceptually grouped into threes. In the unpredictable 3-3-3 condition, the locations of the two extended pauses varied from trial to trial to create 21 different patterns of grouping. These were 3-3-3, 1-6-2, 6-1-2, and all permutations of 2-4-3, 4-4-1, 2-5-2 and 3-5-1. Lists grouped 3-3-3 were presented on trials 2, 8, 13, 20, 26, 30, 33, 39, 43 and 48. The 20 irregular grouping patterns each occurred twice in a random order in the remaining trials.

Procedure. Participants were tested individually in a quiet room in the presence of the experimenter. In the no suppression condition, a trial began with the presentation of the words "Get Ready" in the centre of the computer display followed by a fixation cross. A list of nine spoken digits presented via the headphones then followed. In the ungrouped condition, the interstimulus interval (ISI: offset-to-onset) was 300ms, whilst in the unpredictable 3-3-3 and predictable 3-3-3 conditions the within-group ISI (the interval separating items within groups) was 150ms and the between-group ISI (the interval separating groups) was 750ms. These timings equated list presentation durations in all conditions. At the end of each trial a question mark presented in the central screen position signified the cue to begin recall. Participants subsequently reported the sequence in forward serial order by writing each item in one of nine locations on a response sheet, where each location represented an item in the to-be-recalled list. The recall interval was fifteen seconds in duration and upon completion the next trial commenced automatically. To alert participants to the beginning of the next trial a tone sounded two seconds before the recall interval expired.

The structure of the trials in the suppression condition was the same as outlined above, except that the “Get Ready” signal was replaced with the word “Suppress”. At the onset of this signal participants spoke the word “the” repeatedly at the rate of two utterances per second until the recall cue appeared on screen. The experimenter monitored participants at all times to ensure compliance with the suppression instructions. If participants failed to keep to the rate of two utterances per second they were encouraged to try harder. Both the no suppression and suppression blocks were preceded by three practice trials.

2.1.2 Results

Only data for trials corresponding to those in which 3-3-3 lists were presented in the unpredictable condition were subjected to analysis (viz. trials 2, 8 13, 20, 26, 30, 33, 39, 43, and 48 of each condition). Restricting analysis to these trials ensured that comparisons between conditions were based on equal numbers of observations made at corresponding points. The data were scored using a strict serial recall criterion: an item was only scored as correct if its output serial position was the same as its input serial position.

Accuracy. Figure 4 shows the accuracy serial position curves for each list-type for (a) the no suppression condition and (b) the suppression condition. Considering first the no suppression condition, it can be seen that performance was comparable for predictable and unpredictable 3-3-3 lists and greater than for ungrouped lists. The serial position functions for predictable and unpredictable 3-3-3 lists are characterised by multiple bowing within groups combined with primacy and recency across the list as a whole. By comparison, the serial position function for ungrouped lists shows steeper primacy and recency effects over the list as a whole. The curves for ungrouped lists were also slightly scalloped, suggesting some subjective grouping took place. Performance in the suppression condition was generally poorer, but the pattern was broadly similar. Thus, the recall advantage for predictable and unpredictable 3-3-3 lists persisted, as did the multiple bowing of the serial position curve. Interestingly, recall of unpredictable 3-3-3 lists was actually slightly better than for predictable 3-3-3 lists due to superior recall of items in the first group.

Figure 4 here

A 3 (list-type: ungrouped vs. predictable 3-3-3 vs. unpredictable 3-3-3) × 2 (secondary-task: suppression vs. no suppression) ANOVA revealed a significant main effect of list-type [$F(2, 27) = 9.08$, $p < .01$], a significant main effect of secondary-task [$F(1, 27) = 102.6$, $p < .001$], and a significant list-type × secondary-task interaction [$F(2, 27) = 4.74$, $p < .05$]. The interaction materialized because in the no suppression condition recall of predictable and unpredictable 3-3-3 lists did not differ significantly [$t(18) < 1$], whereas in the suppression condition unpredictable 3-3-3 lists were recalled with greater accuracy than predictable 3-3-3 lists, although this comparison did not reach conventional significance levels [$t(18) = 1.42$, $p = .17$]. In both conditions recall of predictable 3-3-3 lists was better than for ungrouped lists [$t(18) = 2.58$, $p < .05$, and $t(18) = 3.20$, $p < .01$], respectively.

Errors. Errors were classified as omissions (items not recalled) and transpositions (items recalled in the wrong serial position). As expected, omissions were much less frequent overall than transpositions (5% vs 29% of responses), and grouping reduced transpositions (ungrouped 41%, unpredictable grouped 23%, predictable grouped 25%). Suppression reduced performance primarily by increasing transpositions (from 21% to 38%). Transpositions were further classified according to transposition distance, defined as the absolute numerical difference between an item's input and output serial positions. Figure 5 shows the transposition gradients for each list-type, for (a) the no suppression condition and (b) the suppression condition, in terms of the proportion of transpositions as a function of transposition distance. As can be seen, the curves for predictable and unpredictable 3-3-3 lists exhibit interpositions, as reflected by larger amplitude peaks at transposition distance 3 and to a lesser extent distance 6. In contrast transposition gradients for ungrouped lists showed a much smoother decrease with increasing transposition distance. However, it is apparent that the gradients for ungrouped lists exhibit small deviations at transposition distances 3 and to a lesser extent distance 6. Such deviations suggest that, in line with the modest scalloping of the associated serial position curves, some degree of subjective grouping into threes took place in this condition.

Figure 5 here

Table 1 shows the proportions of transpositions within and between groups as a function of list-type and secondary-task. Transpositions between groups are further sub-divided into interpositions and other between-group transpositions. Inspection of this table reveals that predictable and unpredictable 3-3-3 lists were not only associated with larger proportions of interpositions, but also smaller proportions of other transpositions between groups and within groups (with the exception that in the suppression condition the proportion of transpositions within groups was comparable for ungrouped and unpredictable 3-3-3 lists).

Table 1 here

Two 2 (list-type: ungrouped vs. grouped) × 3 (error-type: within-group transpositions vs. interpositions vs. "other" between-group transpositions) × 2 (secondary task) ANOVAs were carried out on the log-odds transformed error proportions. The first analysis compared predictable 3-3-3 and ungrouped lists, whilst the second compared unpredictable 3-3-3 and ungrouped lists. For the first ANOVA, there was a significant main effect of error-type [$F(2, 36) = 3.69$, $p < .05$], and importantly a significant list-type × error-type interaction [$F(2, 36) = 9.82$, $p < .001$], with grouping increasing the proportion of interpositions [$t(9) = 3.10$, $p <.05$], but decreasing the proportions of other between-group transpositions [$t(9) = 1.90$, $p < .10$] and within-group transpositions [$t(9) = 2.30$, $p <.05$]. For the second ANOVA, there was a similar significant list-type × error-type interaction [$F(2, 36) = 17.15$, $p < .001$]. As before, grouping increased the proportion of interpositions [$t(9) = 4.40$, $p < .01$] and decreased the proportion of other between-group transpositions [$t(9) = 5.28$, $p < .001$]. However, grouping did not significantly modify the proportion of within-group transpositions [$t(9) < 1$]. There was also a significant error-type × secondary-task interaction [$F (2, 36) = 3.96$, $p < .05$], which arose because articulatory suppression did not significantly influence the proportions of within-group transpositions [$t(9) <1$], or interpositions [$t(9) = 1.81$], but did significantly increase the proportion of other between-group transpositions [$t(9) = 3.19$, $p <.05$].

2.1.3 Discussion

The main results of the experiment are straightforward. In the no suppression condition, grouping exerted a multiplicity of effects including an elevation in recall accuracy; effects of primacy and recency within groups; a reduction in transpositions between groups; and an increase in interpositions. These effects were present in equal measure for both predictable and unpredictable 3-3-3 lists. Articulatory suppression depressed performance, as expected, but critically the effects of grouping persisted for both predictable and unpredictable 3-3-3 lists, suggesting that the corresponding effects in the absence of suppression were not attributable to participants subjectively rehearsing lists in threes.

That characteristic grouping effects were observed with unpredictably grouped lists suggests that advance knowledge of the grouping structure is not necessary for their manifestation. Furthermore, that recall of predictable 3-3-3 lists was no better than unpredictable 3-3-3 lists—with recall of unpredictable 3-3-3 lists actually being superior in the suppression condition—additionally suggests that foreknowledge of the grouping structure does not modulate the effects of grouping. The present results therefore favour the view that—in the auditory modality at least—temporal grouping effects are largely attributable to a bottom-up mechanism, with little role for top-down driven strategies based on foreknowledge of the grouping structure. We note, however, that top-down subjective grouping strategies do seem to play a minor role in the recall of *ungrouped* auditory lists, as evidenced by modest scalloping of the serial position curve and interposition errors of the type associated with temporally grouped lists.

One limitation of the current experiment is that the results do not preclude the possibility that foreknowledge may be beneficial for other patterns of grouping besides 3-3-3. It is arguably possible that grouping into threes represents a special case, an assumption supported by evidence this is the most beneficial pattern of grouping (Ryan, 1969a; Wickelgren, 1964, 1967). For example, it has been suggested that this is because for lists grouped in threes the positions of items within groups can be specified in terms of unambiguous “beginning”, “middle”, and “end” positional codes (Wickelgren, 1964, 1967). One could argue that these codes are sufficiently unambiguous that they do not benefit from the further elaboration provided by foreknowledge of the list structure. However, for lists containing groups of more than three items such codes are more ambiguous and advance knowledge of the grouping pattern might prove beneficial. We sought to explore this possibility in our next experiment.

2.2 Experiment 2

The first aim of Experiment 2 was to extend the generality of the previous results by examining the effect of foreknowledge of the grouping pattern for a range of irregular patterns besides regular grouping in threes. To accomplish this, we conducted a replication and extension of the study of Ryan (1969a; Experiment 2). As in Experiment 1, we compared serial recall performance for predictably and unpredictably grouped lists in a between-participants design, but this time we dropped the ungrouped list condition. The unpredictable grouping condition was based on the corresponding condition from the previous experiment, except that we increased the number of grouping patterns from 21 to 28, each pattern being presented on an equal number of occasions. As before, the presentation order of the grouping patterns was random, meaning that participants were unable to predict the pattern of grouping on a forthcoming trial. Note that this condition is identical to Ryan’s Experiment 2 in all respects, except that she used only two presentations of each grouping pattern, whereas we increased this number to ten to improve power and permit more detailed analysis of performance for each pattern. The predictable grouping condition—which Ryan did not include—was identical, except that the different patterns of grouping were presented in blocks of 10 trials, so that within each block participants could predict the pattern of grouping on forthcoming trials.

If our previous failure to observe a beneficial effect of foreknowledge of the list structure on the serial recall of lists grouped into threes is attributable to the special status of this grouping pattern then we would expect recall performance for other grouping patterns to be better when the pattern is known in advance. We would further expect serial position curves for the different grouping patterns lists to exhibit more pronounced scalloping when the pattern is predictable. By contrast, if recall performance and the degree of scalloping of serial position curves for different grouping patterns are unaffected by predictability this would be further support for a bottom-up account of grouping in the auditory modality.

The second aim of Experiment 2 was to compare recall performance for different grouping patterns against predictions of the BUMP model. We discuss this second aspect after having first described the main empirical findings.

2.2.1 Method

Participants, Stimuli, & Apparatus. Twenty-eight members of the campus community at the University of York took part in the experiment in exchange for course credit or payment of £6. The apparatus and stimuli were the same as those used for Experiment 1 except that new quasi-random lists of digits were created for each participant subject to the same constraints outlined previously.

Design & Procedure. The experiment had a 2 (list-type: predictably grouped vs. unpredictably grouped) × 28 (grouping pattern) mixed factorial design. List-type was a between participants factor whilst grouping pattern was partly between and partly within (see below). Participants were allocated randomly to the two predictability conditions. The 28 patterns of grouping comprised 3-3-3 and all permutations of 1-1-7, 1-6-2, 1-3-5, 2-2-5, 1-4-4, 2-3-4. In order to keep the experiment to a respectable length whilst still obtaining a sufficient number of observations for each pattern of grouping, half of the participants received one subset of 14 patterns (ensemble A: 1-1-7; 1-5-3; 1-6-2; 1-7-1; 2-2-5; 2-5-2; 2-6-1; 3-1-5; 4-1-4; 4-3-2; 4-4-1; 5-1-3; 5-2-2; 7-1-1), whilst the other half received the other subset (ensemble B: 1-2-6; 1-3-5; 1-4-4; 2-1-6; 2-3-4; 2-4-3; 3-2-4; 3-3-3; 3-4-2; 3-5-1; 4-2-3; 5-3-1; 6-1-2; 6-2-1). There were 10 lists for each pattern of grouping. For unpredictably grouped lists, patterns were assigned to trials in a different random order for each participant, whereas for predictably grouped lists, patterns were presented in 14 blocks of 10 trials, each block involving a single pattern. The order of grouping patterns across blocks for each participant was determined by a Latin Square.

The procedure was identical to that for the no suppression condition of Experiment 1.

2.2.2 Results

Overall accuracy. Figure 6 shows accuracy serial position curves, averaged across the different grouping patterns, for predictably and unpredictably grouped lists. By inspection performance on the two list-types was near identical. Also apparent is that the shape of the serial position curves resembles that expected for ungrouped lists, with a smooth monotonic decrease in recall performance from the first position onwards initially, followed by an upturn in the trend line towards the end of the list. As we will see shortly, the smooth and continuous appearance of these aggregate curves disguises systematic and substantial variability in the shapes of the curves for the different grouping patterns. The overall performance data were subjected to a 2 (list-type: predictably grouped vs. unpredictably grouped) × 9 (serial position) ANOVA, which revealed a significant main effect of serial position, [$F(8, 208) = 104.48$, $p < .001$]. However, crucially, neither the main effect of list-type [$F(1, 26) < 1$], nor the interaction between list-type and serial position [$F(8, 208) < 1$], were significant.

Figure 6 here

Accuracy by grouping patterns. The proportion of correct responses for each of the different groupings for predictably and unpredictably grouped lists is illustrated in Table 2. It is clear that there was considerable heterogeneity in recall performance across the different groupings for both list-types, ranging from 51.4% (unpredictable 1-7-1) to 80.2% (predictable 3-3-3). Crucially, however, for the majority of groupings recall performance was very similar for predictably and unpredictably grouped lists.

Table 2 here

The proportion correct data were subjected to two separate 2 (list-type) × 14 (patterns) ANOVAs, one for ensemble A the other ensemble B. We have already established that overall recall performance for predictably and unpredictably grouped lists did not differ statistically. What is of interest in the present analyses is whether any interactions between list-type and patterns are significant, indicating that, for some groupings at least, recall performance differed significantly between the two list-types. The analyses revealed this was not the case. For ensemble A, there was a significant main effect of patterns [$F(13, 156) = 7.24$, $p < .001$], however, importantly neither the main effect of list-type [$F(1, 12) < 1$], nor the list-type × patterns interaction [$F(13, 156) < 1$], reached significance. For ensemble B, there was once again a significant main effect of patterns [$F(13, 156) = 4.23$, $p < .001$], however, once more, neither the main effect of list-type [$F(1, 12) < 1$], nor the list-type × patterns interaction [$F(13, 156) = 1.19$], was significant.

To further illustrate the close correspondence between performance in the predictable and unpredictable conditions for the different groupings, panel (a) of Figure 7 shows the scatterplot. As can be seen there is a strong positive relationship between performance on the different groupings for the two list-types (r =.721). It is also useful to consider the relationship between performance on the different groupings in the original study of Ryan (1969a) and performance on the same groupings in the present study. Panel (b) of Figure 7 shows a scatterplot of Ryan's data against the present data for predictably grouped lists, whilst panel (c) shows Ryan's data against the present data for unpredictably grouped lists. As can be seen, there are strong positive correlations with Ryan's data in each case (r = .738 for predictably grouped lists, and r = .716 for unpredictably grouped lists). The similarity of the pairwise correlations between these three different data sets is striking.

Figure 7 here

Interestingly, the product of the group sizes is a good predictor of recall performance across the different grouping patterns in each data set (r =.763 for predictably grouped lists, r = .751 for unpredictably grouped lists and r = .731 for Ryan's data). Indeed, the predictive value of the product of the group sizes is at the limit set by the reliability of the behavioural data, as reflected by the correlations shown in Figure 7.

Serial position curves. Figure 8 shows sample serial position curves for four different grouping patterns for predictably and unpredictably grouped lists: 2-6-1; 4-4-1; 2-3-4 and 3-3-3. Curves for all 28 grouping patterns confirm the general picture and can be inspected as supplementary materials (Figures S1-18). In marked departure from the smooth and continuous appearance of the aggregate serial position curves (see Fig. 6), these curves often—but not always—exhibited multiple scalloping consistent with the pattern of grouping. This was the case for both predictably and unpredictably grouped lists, with no evidence that the degree of scalloping was any greater when the grouping pattern was predictable.

Figure 8 here

2.2.3 Discussion

The present results extend the generality of those of the previous experiment by showing—for a variety of different temporal grouping patterns—that foreknowledge of the grouping structure on an upcoming trial does not benefit serial recall. Unpredictably grouped lists were recalled with levels of accuracy indistinguishable from predictably grouped lists, and performance on the different grouping patterns was strongly correlated between the two list-types. The different groupings also exerted similar effects on the serial position curves for the two list-types, the curves for both often exhibiting scalloping consistent with the pattern of grouping. These results are a further indication that the impact of a particular pattern of grouping on serial recall is not modulated by top-down strategies based on knowledge of the grouping structure, and provide—in our view—compelling evidence that in the auditory modality positional information is encoded by a bottom-up mechanism sensitive to grouping structure.

Consistent with Ryan's (1969a) original findings, we observed considerable heterogeneity in recall performance across the different groupings (see also Supplementary Information, Table S1). Indeed, there was a strong positive correlation not only between performance on the different grouping patterns for predictably and unpredictably grouped lists, but also with Ryan's data. Like Ryan, we found that regular grouping into threes was the most effective form of grouping, whilst irregular groupings containing one large and two small groups (e.g., permutations of the patterns 7-1-1 and 6-2-1) were generally the least effective. Ryan suggested that the efficacy of temporal grouping is a function of the regularity of the grouping.

This assertion is supported by our observation that the product of the group sizes is a potent predictor of recall performance (see also Supplementary Information, Table S2) in that this provides a metric for the degree of regularity of the grouping pattern. Small products are associated with irregular groupings where one group is much larger than the others, whereas the largest product and highest level of recall corresponds to equal groups of three. However, while the "product rule" provides an excellent empirical yardstick for predicting the overall effectiveness of different grouping patterns, it does not explain *how* such patterns determine overall recall performance, or have anything to say about more detailed characteristics of recall such as serial position curves and errors. As we shall see shortly, these characteristics are emergent properties of the bottom-up mechanism employed in the BUMP model. For now, we note that the central finding of the current experiment—that predictably and unpredictably grouped lists are recalled with equal efficacy—is consistent with the bottom-up feature of the BUMP model. However, the plausibility of the model would be further bolstered if it could be demonstrated that it can provide a quantitative account of the heterogeneity in recall performance across the different groupings. We turn to this issue in next section.

3. Simulations of grouping

The detailed implementation of the BUMP model is described in the Appendix. For simplicity, simulations used a minimal implementation of competitive queuing to assess whether the key empirical phenomena can be explained solely in terms of properties of the context signal. Thus, associations between items and temporal context are stored simply as the mean activation of the context vector (i.e. the activity of the set of idealized "neurons" in the model) during the presentation of each item. To determine an item's activation at retrieval we replay the context signal and compare each item's stored context vector with the current state of the context signal using a Euclidean distance metric: the closer the stored context vector is to the current context, the more active the item becomes. As each item's activation peaks, it competes with the other items to determine which is recalled (we simulate a large number of competitions in which Gaussian noise is added to each item's activation, leading to occasional errors). Once output, the item's activation is suppressed which is normally sufficient to prevent it from being reselected—the degree of suppression decays exponentially over time. We analyze the winning item at each serial position.

Briefly we simulate the retrieval of a temporally structured list by:

i) Creating an input signal based on amplitude modulations associated with the presentation of each item. The timing of the onsets of the items follows that used by Ryan (1969a). The duration of the items is not recorded in the original paper, but in the simulations it is held constant (0.3s). (Examples of the input signal used for ungrouped, and 3-3-3 and 2-6-1 grouped lists were shown in Figure 3).

ii) The context signal generated by each input signal is calculated. The context signal is the time-varying response of a population of cells, which act as temporal filters with different AM tunings (scales), spanning the range of frequencies encountered in the task. In the simulations reported here the population comprised 15 log-spaced frequencies ranging from 0.10Hz to 1.28Hz (see Appendix). The precise details are not critical to the model's key predictions, though the overall span and log-spacing may have important effects, as we discuss later. Each frequency is represented by two phase-shifted filters (equivalent to a complex quadrature filter). Examples of the evolution of phase and amplitude responses to ungrouped, and 3-3-3 and 2-6-1 grouped lists are shown in Figure 3.

iii) Each item is associated with the state of the context signal during its presentation, which is stored.

iv) At retrieval the context signal is replayed. Items to be recalled are activated according to the similarity between the current state of the context signal and the stored association. Items compete to be output, with the most active item being selected at the time points where the similarity of the current and stored states of the context signal is greatest. However, activation also includes Gaussian noise, which means that the correct item will not always be selected. After selection, the activation of the item is subject to inhibition which typically prevents it being reselected in immediately following serial positions.

v) The order of the retrieved items is compared to the correct order, and responses at each serial position are recorded as correct or incorrect, with transpositions of ±1, ±2, ±3 also being counted separately.

We simulated 100,000 retrieval attempts for each grouping pattern considered by Ryan (1969a).

The model's parameters (described in the Appendix) were selected to give overall levels of performance approximately in line with experimental observation (items correct, lists correct, serial position curves for ungrouped lists) but have not been fitted to data from grouping studies.

3.1 Results

We focus on comparing simulations of the recall of irregularly grouped lists with data from Ryan (1969a) and from Experiment 2. Before doing so we note that the absence of any effect of the predictability manipulation in Experiment 2 allows us to simplify reporting by pooling data over the predictable and unpredictable grouping conditions.

Error rates for different irregular grouping patterns. As can be seen in Figure 3, the responses of low frequency oscillators sensitive to overall list position are qualitatively similar across grouping conditions. Grouping typically leads to the recruitment of extra oscillators and improves the overall level of recall relative to ungrouped presentation, differences between grouping conditions being largely accounted for through the distinctive pattern of recruitment of higher frequency oscillators they engage. As should be clear, however, the degree of improvement will depend crucially on the frequencies present in the grouping pattern and the local regularity of delays. In general, patterns with similar group sizes, and therefore a more homogeneous rhythm, will recruit subsets of oscillators more strongly than patterns with disparate group sizes. However, the important question is whether the model can reproduce human data at a more detailed level of description.

Figure 9 shows scatterplots of observed vs predicted percent correct recall for the 28 grouping patterns studied in Ryan (1969a) and Experiment 2. The correlations between the simulated and observed data were r = 0.768 for Ryan's data and r = 0.737 for the data from Experiment 2 (where predictable and unpredictable conditions have been pooled). As the correlation between the two sets of empirical data was only slightly higher at r = 0.784, the simulations can be regarded as an excellent fit. The simulated data also correlate highly with the product of the group sizes, r = 0.779, which we noted previously is a simple yardstick of the regularity in the grouping pattern. Full data on performance in each grouping condition from both Ryan (1969a), Experiment 2 and our simulations are presented as supplementary data in Table S1. Inter-correlations between these measures are tabulated as Table S2.

Figure 9 here

Serial position curves. The next question concerns the ability of the simulations to predict performance for different patterns of grouping at the finer grain of serial position effects and patterns of errors. Figure 10 shows simulated and observed serial position curves for correct responses and three distances of transposition errors (±1, ±2, ±3) for the four grouping patterns illustrated in Figure 8 (with observed data pooled over predictability conditions to maximize reliability). The patterns sample a range of different products and orders of group sizes. The 24 other patterns can be inspected as ancillary material (Figures S1-18) and confirm the general picture described here.

Figure 10 here

As noted previously, the human data vary markedly with the pattern of grouping, showing multiple scalloping in the curves for correct recall in many cases, a locality constraint on transpositions, and a tendency for adjacent transpositions not to cross group boundaries. The simulated serial position curves show a high degree of resemblance to the human data. Thus in all cases there is primacy and recency for the list as a whole and clear bowing within groups for groups of size four or more. There is also a broad degree of correspondence between the observed and predicted distributions of different types of order error for the various grouping patterns in that simulated transposition errors obey the locality constraint but adjacent transpositions tend not to cross group boundaries. There are however some discrepancies, perhaps the most noticeable being that the simulations generate weaker recency than the human data, slightly under predicting recall of the final item and slightly over predicting performance on the penultimate item. Perhaps relatedly, the multiple bowing of the serial position curve is less pronounced for smaller groups, notably 3-3-3 lists.

As we have seen, according to the model the effect of regular grouping is to recruit extra oscillators at frequencies close to the group presentation rate whose phase changes progressively from the beginning to the end of each *group*. The effect of this is to decrease the similarity of the context signals associated with adjacent items relative to ungrouped presentation, resulting in fewer adjacent transpositions and an increase in proportion correct. There is however a cost as the phase information associated with items in corresponding positions in different groups is now more similar, resulting in a very clear increase in interposition errors (±3 transpositions). Thus, in summary, the simulations reproduce the overall pattern of human data with remarkable accuracy, although there are some minor systematic discrepancies which we discuss below.

3.2 Discussion

The BUMP model provides a detailed implemented account of the mechanisms underlying grouping effects and explains how, by exploiting temporal structure in the input sequence, the capacity and fidelity of auditory-verbal memory can be extended. Experimental data clearly show that serial recall can benefit from such structure in the input, even when it is unpredictable and irregular. This can only be possible within a competitive queuing mechanism if the context signal is sensitive to such structure. In implementing such a general context signal (i.e., one not specialised for a particular pre-specified temporal structure) we extended the competitive queuing framework to these circumstances. We find that many otherwise unexplained features of serial recall can be understood in terms of the similarity of states of the context signal which are associated with different items. As in all competitive queuing models, items associated with similar states tend to transpose with one another. In the BUMP model, this is neither a monotonic function of time nor serial position, nor “hard-wired” for a particular expected structure. It is an emergent property shaped by the temporal structure of the input and the resonance it produces in an array of oscillators tuned to local fluctuations on a range of timescales. This population produces a coherent response, especially when driven by regularly timed inputs; grouping typically serves to recruit a wider range of oscillators, with some grouping patterns producing stronger and more coherent activity than others.

When overall performance on different grouping patterns is compared, the pattern of results closely matches those seen in experiments, with the correlation between data and model being as strong as the consistency between experimental studies. The model also reproduces familiar features of serial memory: a bowed serial position curve showing primacy and recency, the tendency for local transpositions and interposition errors in regularly grouped sequences. Detailed qualitative features of irregularly grouped list recall are also reflected in the simulations; in particular we see that the serial position curve within longer groups is clearly bowed so that the shape of the serial position curve is sensitive to the grouping structure.

The presence of one or two systematic discrepancies between the simulation results and empirical findings indicates scope for refining the model. In the simulations we do not see clear bowing of the serial position curve for shorter groups whereas in the empirical data this is often fairly pronounced leading to a scalloped serial position curve, especially in regularly grouped lists, even though we see pronounced multiple bowing in irregular sequences. The model has somewhat reduced recency compared with the experimental data, and its performance on regularly grouped (3-3-3) lists does not benefit as much as that of human participants. With its current parameters, the model appears to be slightly too sensitive to regular grouping with a large proportion of oscillators resonating with grouping frequency for regularly grouped lists, leading in turn to a high proportion of interposition errors at all serial positions which may contribute to some of the differences we see.

We have not experimented systematically with the parameters governing the range and spacing of tunings of neurons in our population but they are likely to have some subtle effects on the overall behavior of the model. The similarity metric we use to determine the activation of items at retrieval (and hence their propensity to transpose with one another) weights each oscillator equally. If the population has more low-frequency oscillators (at or near the list presentation rate) then it will be more prone to local transpositions and correspondingly less prone to interposition errors. If the population has more mid-frequency oscillators (corresponding to the temporal scale of groups encountered in the input) then the overall response of the population will tend to resonate with the group (especially for regularly grouped lists) and interposition errors will be more likely. At even higher frequencies, the oscillators will go through one cycle for each item encountered. These oscillators do not contribute to the encoding of serial order at the timescales involved in serial recall experiments, though they might be useful for encoding phonological sequences (e.g., nonwords); in serial recall high frequency oscillators make the context-signal associated with consecutive items more similar and are therefore likely to oppose the effects of grouping.

4 General Discussion

In the current study our aims have been to better understand the relationship between rhythm and memory for spoken sequences and to develop an account of a mechanism that can explain this relationship. Rhythm has been implicated in the general problem of serial order since it was set out by Lashley (1951), and temporal grouping has been known to have large effects on the serial recall of verbal materials since the late 1960s (Ryan, 1969a, 1969b). However, despite progress in computational modelling of immediate serial recall, the effects of irregular and unpredictable temporal grouping have been largely unaddressed. This is a significant problem since the timing of natural utterances is neither regular nor predictable, and as we showed in Experiments 1 and 2, the effects observed by Ryan do not depend on predictability and are highly and reproducibly sensitive to different patterns of irregularity. We argue that these data are strongly suggestive of a bottom-up mechanism for the encoding of serial order, meaning one that is sensitive to the timing of spoken items as they occur.

We described a multi-oscillator filter mechanism (BUMP) that responds to local changes in the speech signal on a range of timescales, and we showed that it generated predictions for the recall of sequences grouped in a variety of different ways that were a good match to human data. The specification of a plausible stimulus-driven mechanism for determining the context signal makes the BUMP model a significant advance on existing context signal models of short-term memory, and, we suggest, a platform for future investigation and development. In the remaining discussion we consider the empirical findings in more detail before moving on to discussing the BUMP model, implications for current models of serial order in STM and finally broader implications for serial order in language processing more generally.

4.1 Experimental findings

We examined the role of encoding strategies by assessing whether the ability to recall a sequence depended on whether its grouping pattern could be predicted in advance. We reasoned that if such strategies do play a significant role, there should be stronger effects of grouping when the pattern is known in advance. Experiment 1 compared recall of regularly grouped (3-3-3) sequences when they occurred on every trial and when they were randomly embedded amongst trials involving other grouping patterns. Grouping led to the same improvement in recall in each case when compared with an ungrouped control condition. The grouping effect remained when rehearsal was disrupted by articulatory suppression and there was again no benefit of knowing the grouping pattern in advance. These results are consistent with the suggestion that strategies such as rehearsal are not an important mediator of grouping effects for auditory sequences (Frankish, 1985; Hitch et al 1996, Ryan, 1969b). Experiment 2 compared recall of a total of 28 irregular grouping patterns. Results showed large variation in overall levels of recall, serial position curves, and error patterns for different grouping patterns but no effect of whether or not the grouping pattern could be anticipated in advance of sequence presentation. Thus it was not simply that the predictability of the grouping pattern had a null effect, which by itself would be difficult to interpret, but predictability had no effect alongside large and systematic variation in recall associated with different grouping patterns. These findings suggest that for auditory-verbal sequences, sensitivity to grouping is primarily dependent on bottom-up processes. The importance of bottom up processes does not rule out any contribution of top-down strategies, and indeed it is clear that It is important to suffix this conclusion with one qualifying note, namely that subjective grouping strategies do nevertheless seem to play a minor role, for example, in memory for ungrouped auditory-verbal sequences. Thus in Experiment 1, serial position curves for ungrouped sequences showed modest scalloping suggesting a spontaneous tendency to impose a 3-3-3 pattern on the input sequence, consistent with other data (Farrell & Lelievre, 2009; Henson, 1996, Madigan 1980). Broadly however, our results indicate that the critical variations in accuracy and error distributions for irregularly timed sequences are not explicable in terms of such strategies.

Experiment 2 was also important in showing that Ryan's (1969a) results for irregular grouping patterns were not only highly replicable, but also consistent with a rule of thumb whereby sequence memorability varies as the product of the group sizes. This empirical relationship is of more than passing significance as it places the classic finding that grouping in threes is optimal on a continuum, in contrast to suggestions that the number three represents a special case (e.g., Wickelgren, 1964, 1967). We suggested that the product of group sizes is important because it reflects the regularity of the grouping pattern, which in turn determines the overall distinctiveness of the context signal in BUMP. It remains to be seen whether the product rule applies more generally and, for example, extends to different numbers of groups and different sequence lengths.

4.2 Evaluation of the BUMP model

The two key features of the model are the assumptions (1) that the serial order of items in a sequence is encoded by a bottom-up analysis of the speech input, and (2) that this analysis is achieved by means of a set of filters that respond to local changes in the magnitude and phase of its amplitude envelope at different frequencies. As described above, the first assumption is broadly supported by the present results. The principal support for the second assumption is the BUMP model's capacity to predict overall percent correct for different irregular grouping patterns.

Performance varied markedly with grouping pattern in Experiment 2, and the model's ability to predict this variation was as good as the consistency between the observed data and the previous data of Ryan (1969a). Our simulations also capture key qualitative properties of the distinctive serial position curves and distribution of local transposition and interposition errors that are characteristic of different grouping patterns. Critically, these are emergent properties of the BUMP mechanism; no attempt was made to fit the model to the data beyond initial choice of parameters to reproduce appropriate levels of performance for ungrouped lists, as described earlier. We consider that the mechanistic explanation of these otherwise unexplained phenomena are persuasive evidence in favour of the involvement of a BUMP-like mechanism in serial memory for spoken sequences.

We note, however, some systematic discrepancies between model and data. For example, it over-predicted recall for the least regular 1-1-7, 1-7-1, 7-1-1 patterns and the most regular 3-3-3 pattern. Also, although simulated serial position curves for irregularly grouped lists showed multiple bowing, with order errors obeying the locality constraint but not crossing group boundaries, broadly consistent with the human data, the model tended to predict too little recency, insufficient multiple bowing and too many interposition errors for 3-3-3 lists. As discussed earlier, many of these discrepancies could be addressed by modifying the distribution of filters, which for simplicity are equally spaced (in terms of log frequency) and equally weighted in the present model. Changing these parameters would fine-tune the balance of sensitivities to groups of different sizes and improve the fit to human data. However, they would not address the problem of under-predicting recency. This may indicate a structural limitation of context signal models as Burgess and Hitch (1999) found it necessary to specify an extra process whereby the final item has special status to achieve sufficient recency for auditory sequences.

4.3 Relationship of BUMP to other models

We introduced BUMP as a new member of the family of context signal models for serial order in short-term memory (Burgess & Hitch, 1992; 1999; 2006; Brown et al., 2000; Farrell, 2012; Page & Norris, 1998), going beyond them by specifying a mechanism for generating the context signal from the perceptual input. BUMP shares with other context signal models the core serial ordering principles of competitive queuing among simultaneously active elements followed by suppression of the winning element (see Hurlstone, et al., 2014).

Historically, BUMP derives from an early attempt to specify an input-driven, speech-specific, context signal in Hartley and Houghton's (1996) model of short-term memory for unfamiliar phonological sequences. A key feature of this model was that the phonological input drove the context signal through a cyclical pattern whereby similar states are entered when phonemes occupy corresponding positions in adjacent syllables. The cyclical element of the context signal was required to account for the human tendency, in nonword repetition (and in spontaneous speech), for between-syllable phoneme transpositions to involve phonemes that occupy corresponding within-syllable positions, analogous to interposition errors in immediate serial recall of temporally grouped lists.

The use of quadrature filters to generate a cyclical context signal was discussed by Hartley (1996), and partially implemented in the model proposed by Henson and Burgess (1997), which, like BUMP, assumes that pairs of filters tuned to different frequencies compete to represent the input. However, unlike BUMP, Henson and Burgess did not specify an on-line, stimulus-driven mechanism. Reflecting this limitation, the grouping pattern had to be known in advance of sequence presentation in order for the model to generate the appropriate context signal. Thus, the BUMP model represents a significant extension of Henson and Burgess (1997).

A particularly close neighbor is OSCAR (Brown et al., 2000), which also assumes a multi-oscillator context signal. However, the crucial difference between OSCAR and the BUMP model is that in the former the oscillators are free-running and the mechanism whereby appropriate oscillators are recruited for any sequence is not specified, as Brown et al. (2000) themselves noted. Other close neighbors are the context signal models of Burgess and Hitch (1999; 2006), and Farrell (2012). These models assume a central serial ordering mechanism that is not tied to any specific input modality. One way of viewing BUMP is as a front-end to such models that enables them to deal with spoken inputs, which in turn highlights the importance of specifying corresponding front-ends for other modalities.

An interesting comparison in the literature concerns the distinction between “time-based” and “event-based” models of serial order. In the former, the context signal changes as a function of absolute time, whereas in the latter, the context signal changes only when a new event (e.g., a new item) is experienced. Ng and Maybery (2002, 2005) contrasted time-based and event-based accounts of temporal grouping effects and found support for the latter, in that interposition errors reflected similar relative positions of items within groups rather than similar times of presentation. On the basis of this result Ng and Maybery (2002, 2005) argued against oscillator models, which they regarded as falling into the time-based category. However the BUMP model has characteristics of both time-based and event-based accounts and can be considered a hybrid in this regard. Its time-based characteristics arise from the continuous change in the context signal. Its event-based characteristics arise from its sensitivity to *local* changes in amplitude of the speech signal. Although data from these studies falls outside the scope of the current study, we note that much of the earlier debate around these issues hinges on what may prove to be a false dichotomy.

4.4 Scope of the BUMP model

At this point, we should perhaps make it explicit that we do not propose the BUMP model, or indeed multiple oscillator systems more generally as a panacea. Indeed we have deliberately limited the scope of the model in order to focus on the explanatory power of the emergent properties of its context signal to address otherwise unexplained effects of irregular and unpredictable timing on memory for auditory-verbal sequences. Many other features of auditory-verbal short-term memory are not addressed. In broad terms these fall into two clusters: effects of materials and effects associated with executive control processes. Here we identify the main omissions and comment on how these phenomena might be reconciled with a BUMP-like account of serial order, starting with effects of materials.

In presenting BUMP as a solution to the “front-end problem”, we purposefully did not attempt to model the way phonological variables influence verbal short-term memory, in particular the signature effects of phonemic similarity and word length and their interactions with articulatory suppression. Given that these effects are empirically separable from the effects of temporal grouping (Hitch et al., 1996) and are reasonably well explained by existing context signal models in terms of item representations (see e.g. Burgess & Hitch, 1999; 2006; Brown et al., 2000), we would envisage a full elaboration of BUMP to include phonological representations of items in a broadly similar way.

Another deliberate omission from the BUMP model for the purpose of simplicity was effects of prior learning in immediate recall, as reflected in well-established effects of variables such as lexicality (Hulme, Maugham & Brown, 1991) and wordlikeness (Gathercole, Willis, Emslie & Baddeley, 1991). We note that such effects have been simulated in context signal models by including a learning or long-term memory component (e.g. Burgess & Hitch, 1999;

2006), and thus are not fundamentally incompatible with BUMP. Related to this, BUMP was not designed to address grouping of inputs on the basis of familiar patterns or chunks (Miller, 1956), which we argue involves fundamentally different knowledge-driven processes from those underpinning grouping by perceptual features alone.

Other forms of long-term knowledge affect memory for linguistic stimuli. We note in particular the importance of prior constraints such as phonotactics, syntax and semantics on the processing of serial order in language. Constraints of this sort were referred to generically by Lashley (1951) as order schemata. To the extent that they are unrelated to the perceptual qualities of speech that drive the oscillators in the BUMP model, they clearly fall beyond its scope. We therefore propose BUMP as a key component of the serial ordering mechanism for language but most certainly not the whole story. In more precise terms, we propose BUMP as specifying the stimulus-driven ordering mechanism with which order schemata must interact to account for serial order processing in speech perception and production more generally.

We turn next to the issue of executive control processes. Although we have presented strong evidence that top-down processes play little or no part in the ability to recall auditory-verbal sequences presented in various patterns of temporal grouping, this is of course not to deny the importance of top-down influences in short-term memory more generally. One particularly closely related example is that of subjective grouping, whereby a grouping pattern can be subjectively imposed on a temporally ungrouped input. Our own data illustrated this in Experiment 1 where participants showed signs of spontaneously grouping temporally ungrouped sequences. Spontaneous grouping is a well-established phenomenon (Farrell & Lelievre, 2009; Henson, 1996; Madigan, 1980). Furthermore, it has also been shown that people can subjectively group ostensibly ungrouped sequences in various different ways when instructed to do so

(Wickelgren, 1964, 1967). It is important to be clear that these subjective grouping effects do not challenge our central claim that top-down processes are unimportant in accounting for memory for temporally grouped auditory-verbal sequences. However, on grounds of parsimony alone one would wish to see them incorporated in a common theoretical account. One way this might be achieved, albeit speculative, is by postulating that subjective grouping reflects an internal, top-down input to the BUMP mechanism that is effective when the external stimulus provides a weak signal. Thus, for an ungrouped list, oscillators at rates corresponding to subjective groups could be given a strategic boost during encoding and retrieval. However, any such internal boost would be ineffective when oscillators corresponding to objective groups were being strongly driven by the external input. More generally, we can imagine top-down modulation of BUMP oscillators could be used to optimize the representation of serial order at different timescales. For example, in the current digit span like tasks, the critical serial structure operates at the level of lexical items, whereas in nonword repetition, finer temporal structure at the scale of phoneme sequencing is more critical. These different tasks might depend on the activity of different subsets of the oscillator population which could be strategically modulated to optimize performance.

Other, related grouping phenomena that we have not addressed in the current study are the non-temporal grouping effects associated with changes in perceptual features such as intonation (Frankish, 1995) and spatial location (Parmentier & Maybery, 2008). Such effects are consistent with the assumption that grouping reflects bottom-up processes operating on the speech input, but the question is whether the BUMP mechanism is capable of handling them. Again we speculate that these non-temporal grouping effects might be reconciled with a BUMP-like mechanism by adding top-down attentional filters which could modulate the degree to which oscillators are engaged by different auditory streams, providing an additional grouping cue.

If top-down input can modulate the degree to which auditory stimuli engage the oscillators, we suspect that such filters will not entirely abolish modulation by unattended streams; we know that unattended speech disrupts immediate serial recall (Salamé & Baddeley, 1982; Jones & Macken, 1993), and in the BUMP model this can be conveniently explained in terms of interference in the entrainment of oscillators by amplitude modulation in the unattended stream. Indeed it is a clear prediction of the BUMP model that (in the absence of a wholly effective attentional filter) AM stimuli at or near speech production rates should be particularly potent disruptors of immediate serial memory, and it should be possible to specify stimuli that are optimal disruptors or which drive specific patterns of serial order error.

A final set of questions concern whether BUMP can address grouping effects for non-verbal sequences. We note firstly that serial recall of sequences in the spatial and visual domains appears to involve competitive queuing in which items are simultaneously active in parallel and the strongest chosen for output (Hurlstone et al., 2014), and that this is broadly consistent with, but not unique to BUMP. However, given that BUMP is specialized for processing speech input, we would not expect to see detailed correspondences with other modalities. Interestingly, there are indications that empirical effects of temporal grouping in the visual and spatial domains differ slightly from those for spoken sequences, and are influenced by domain-specific perceptual features (Hurlstone et al., 2014). We favor the general hypothesis that processing serial order involves a common general principle of competitive queuing across domains driven by domain-specific mechanisms for analyzing the perceptual input, but another possibility is that a common context signal (such as BUMP) can be repurposed for different tasks by the application of top-down modulation.

4.5 Broader implications

We now turn to the important question of whether the BUMP mechanism has relevance beyond short-term memory. We referred earlier to the empirical evidence for a link between phonological short-term memory and vocabulary acquisition (Baddeley, Gathercole & Papagno, 1998) and the idea that it is best viewed as part of the language system (Allen & Hulme, 2006; Martin & Saffran, 1997; Monsell, 1987). The link with vocabulary acquisition has stimulated proponents of computational models of short-term memory to regard them as integrative accounts that embrace vocabulary acquisition as well as immediate serial recall (e.g., Hartley & Houghton, 1996; Page & Norris, 2009; Gupta & MacWhinney, 1997; see also Burgess & Hitch, 2006). These authors argue that the representation of serial order in immediate memory is crucial for forming long-term representations of familiar words as sequences of phonemes. There are also grounds for considering whether a BUMP-like mechanism supports speech production. For example, similarities between error patterns in speech production and immediate serial recall of verbal sequences have been interpreted as suggesting a common serial ordering process (Page, Madge, Cumming, & Norris, 2007). Work by Vousden, Brown and Harley (2000) is particularly interesting in that they show how a multi-oscillator based process similar to that proposed by Brown et al. (2000) might explain a range of phonological constraints on speech errors – in which reordered segments typically retain their syllable positions. Their model is broadly compatible with the BUMP mechanism, but lacked a mechanism to explain how the oscillator phases might be entrained to stimuli so that appropriately structured representations could be learned in the context of natural, continuous speech. However, preliminary work by Hartley

(2002) shows that the coherent response of a BUMP-like mechanism tracks the production of syllables in connected speech such that the overall phase of the population response represents the current syllable position. BUMP therefore has the potential to provide a link underpinning the control of serial order in nonword repetition (Hartley & Houghton, 1996; Treiman & Danis, 1998), vocabulary acquisition (Gupta & MacWhinney, 1997) and speech production (Vousden et al., 2000).

A further opportunity for broader theoretical integration is with models of speech perception. It is interesting to note evidence that speech analysis involves multiple oscillators that sample the amplitude envelope at different time-scales to allow segmentation at different grain sizes corresponding to phonemes, syllables, etc. (Luo & Poeppel, 2007; Poeppel, Idsardi, & van Wassenhove, 2008). We note also recent MEG data demonstrating that speech entrains the phase of low frequency oscillations in auditory cortex and that edges in the speech envelope enhance this entrainment by resetting the phase (Gross, Hoogenboom, Thut, Schyns, Panzeri, Belin & Garrod, 2013). The degree of similarity of this proposal with BUMP is remarkable in that the oscillators analyzing the speech input operate over a similar range and are sensitive to local changes in the phase of the input signal. It suggests that the same processes might be involved in both perceptual segmentation (in terms of the phase of BUMP oscillators; – see Hartley, 2002) and the encoding of serially-ordered representations of speech.

A connection between speech perception, rhythm and language learning through a common BUMP-like mechanism might have wider implications for understanding specific disorders of memory and language. We note for example, that deficits in neural oscillators of the type proposed by Poeppel and colleagues have recently been proposed as an integrative theoretical framework for understanding speech perception difficulties associated with the developmental disorder of dyslexia (Goswami, 2011; Power, Mead, Barnes & Goswami, 2012). With this in mind, the BUMP mechanism could offer an explanatory link between observations of perceptual and timing deficits on the one hand and serial memory and phonological learning deficits on the other. In this view some developmental disorders of language learning would be expected to derive from an underlying problem in the processing of amplitude modulation: for example, individuals with more broadly-tuned oscillators are likely to be more prone to ordering errors in phonological memory, less sensitive to rhythmic structure and less well able to exploit temporal cues to forming new and robust representations of verbal sequences. Our current simulations suggest that the pattern of grouping effects seen in auditory-verbal short-term memory may provide a detailed behavioral signature of the operation of oscillatory coding processes. Investigation of these phenomena in developmental disorders could provide a useful test of Goswami's (2011) proposals and the suggested link between mechanisms of short-term memory and language development.

4.6 Conclusion

We have shown that temporal grouping effects in auditory-verbal short-term memory are consistent with the emergent properties of the BUMP model’s, bottom-up multi-oscillator mechanism that associates items with the local changes in the phase and amplitude of the speech signal. We have described how this theoretical account is consistent with, but goes substantially beyond current competitive queuing models of short-term memory by providing a plausible account of how the context signal in such models is derived from the perceptual input. We tentatively propose BUMP as a plausible instantiation of Lashley’s (1951) original speculation that serial behavior, at least in the domain of language, is mediated by a rhythmic representational system. Although the BUMP model is presently limited to short-term memory for auditory-verbal inputs, the observation of similarities in serial order phenomena across domains leads us to hypothesize that the multi-oscillator system acts as a common ordering mechanism in vocabulary acquisition, speech perception and speech production.

Acknowledgements

Portions of this work were funded by an Economic and Social Research Council (United Kingdom) 1+3 studentship awarded to Mark Hurlstone, and by a small grant awarded to Graham Hitch, Mark Hurlstone, and Tom Hartley by the Department of Psychology at the University of York, England. We are grateful to Alan Baddeley for useful discussions of ideas presented here.

References

Allen, R., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language, 55*, 64-88.

Baddeley, A. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D. (2007). *Working memory, thought and action*. Oxford: Oxford University Press.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*, 575-589.

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

Broadbent, D. (1975). The magic number seven after 15 years. In A. Kennedy & A. Wilkes (Eds.), *Studies in long term memory* (pp. 3-18). New York: Wiley.

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

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

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.

Chi, M. T. H. (1976). Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition, 4,* 559-572.

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

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*, 123-147.

Farrell, S. (2012). Temporal clustering and sequencing in short-term and episodic memory. *Psychological Review, 119*, 223-271.

Farrell, S., & Lelievre, A. (2009). End anchoring in short-term order memory. *Journal of Memory and Language, 60*, 209-227.

Farrell, S., & Lewandowsky, S. (2002). An endogenous model of ordering in serial recall. *Psychonomic Bulletin & Review, 9*, 59-60.

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. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24*, 200-209.

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 469-479.

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

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on childrens repetition of nonwords. *Applied Psycholinguistics, 12*, 349-367.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia *Trends in Cognitive Science, 15*, 3-10.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P. & Garrod, S. (2013). Speech rhtyhms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology, 11*, 1 – 14.

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics, 31*, 423 – 445.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hartley, T. (1996) The role of syllable structure in verbal short-term memory. PhD Thesis, University of London.

Hartley, T. (2002). Syllabic Phase: A bottom-up representation of the structure of speech In J. Bullinaria & W. Lowe (Eds.), *7th Neural Computation and Psychology Workshop* Singapore: World Scientific.

Hartley, T., & Houghton, G. (1996). A linguistically constrained model of short-term memory for nonwords. *Journal of Memory and Language, 35*, 1-31.

Henson, R. N. A. (1996). *Short-term memory for serial order. Unpublished doctoral thesis*. Unpublished doctoral thesis, Cambridge University Cambridge UK.

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

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

Henson, R.N. & Burgess, N. (1997). Representations of serial order. In J.A. Bullinaria, D.W. Glasspool,

& G. Houghton (Eds.), *4th Neural Computation and Psychology Workshop*. London: Springer.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology, 49A (1)*, 80-115.

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*, 140-158.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish & M. Zock (Eds.), *Current Research in Natural Language Generation* (pp. 287-319). London: Academic Press.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words - evidence for a long-term-memory contribution to short-term-memory span. *Journal of Memory and Language, 30*, 685-701.

Hurlstone, M., Hitch, G. J., & Baddeley, A. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology, 60*, 505-511.

Jones, D.M. & Macken, W.J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition, 19*, 369-381.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium*. New York: John Wiley.

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

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *The Psychology of Learning and Motivation, 49*, 1-48.

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

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron, 54*, 1001-1010.

McNicol, D. & Heathcote, A. (1986). Representation of order information: An analysis of grouping effects in short-term memory. *Journal of Experimental Psychology: General, 115,* 76 – 95.

MacKay, D. G. (1970). Spoonerisms: The structure of errors in the serial order of speech. *Neuropsychologia, 8*, 323–350.

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

Martin, R. & Saffran, E.M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for a common system. *Cognitive Neuropsychology, 14*, 641-682.

Maybery, M., 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.

Miller, G. A. (1956). The magical number 7, plus or minus 2 - some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Monsell, S. (1987). On the relation between lexical input and output pathways for speech. In A. Allport, D.G. MacKay, W. Prinz, E. Scheerer (Eds.), *Language perception and production: Relationships between listening, speaking, reading and writing* (pp. 273-311). London: Academic Press.

Murdock, B. B. (1995). Developing TODAM: Three models for serial order information. *Memory &*

Cognition, 23, 631-645.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21*, 263-276.

Ng, L. H., & Maybery, M. T. (2002). Grouping in verbal short-term memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55A*, 391-424.

Ng, 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*, 175-181.

Page, M. P. A., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language, 5*., 49-64.

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

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of The Royal Society B, 364*, 3737-3753.

Parmentier, F. B. R., & Maybery, M. T. (2008). Equivalent effects of grouping by time, voice and location on response timing in verbal serial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1349-1355.

Poeppel, D., Idsardi, W.J., & van Wasssenhove , V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B, 363*, 1071-1086.

Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Front Psychol, 3*, 216.

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

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

Salamé & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150 – 164.

Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology - Learning Memory and Cognition, 14*, 145-152.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology, 41*.

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

Wickelgren, W. A. (1967). Rehearsal grouping and hierarchical organisation, serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology, 19*, 97-102.

Appendix

As described in the body of the text, the BUMP model describes a context signal sensitive to the timing of items based on the envelope of speech.

Input signal

Each item is modelled as a triangular pulse in amplitude, with the peak halfway between onset and offset of the stimulus. Input signals for lists with different grouping structures are modelled by varying the item onset times in line with the experimental timings, durations are held constant (0.3s).

Context signal

The context signal is based on the activity of a population of cells with *n* different amplitude modulation (AM) tunings ranging from f_b to $f_b\lambda^{n-1}$:

$$f_j = f_b\lambda^{j-1} \qquad \text{Eq. (1)}$$

For j=1 to n, where f_b the base frequency (lowest rate to which cells will be tuned) and λ is a spacing factor. These are parameters of the model which are set such that the range of frequencies encompasses the range of relevant frequencies in the task i.e., f_b<1/(list duration), $f_b\lambda^{n-1} \geq 1/(group/item\ duration)$; λ must be greater than 1 and is typically less than 2.

Pairs of cells in the context signal are modelled as complex quadrature filters, with sinusoidally modulated Gaussian form:

$$F_j(x) = k_j\left(\cos(2\pi f_j x)e^{-(x^2/\sigma_j{}^2)} + i\sin(2\pi f_j x)\, e^{-(x^2/\sigma_j{}^2)}\right) \qquad \text{Eq. (2)}$$

Where σ_j is a constant governing the width of the filter, which is inversely proportional to f_j so that $\sigma_j = \sigma_b/\lambda^{j-1}$ where σ_b is the width of the Gaussian used for the base frequency. This ensures that each F_j encompasses the same number of cycles regardless of its frequency tuning f_j. The range $-\sigma_j$ to σ_j always contains the same number of cycles. The normalising factor k_j is inversely proportional to the absolute area under F_j.

The output of the filter R_j at time t is given by is the convolution of F_j with the AM input signal $h(t)$.

$$R_j(t) = \int F_j(u)h(t+u)du \qquad \text{Eq. (3)}$$

In practice this is implemented numerically using the conv function in MATLAB and FIR filters for F_j where the width of the FIR filter window is $6/f_b$. R is calculated in time steps of 0.01s in all our simulations.

Memory encoding

Item-context associations (M_{ij}) are stored as values of each of the n complex elements the time-varying filter output R_j which obtain when a particular item i is being presented. As the items are extended in time, it is necessary to use an average value, and this is weighted according to the concurrent input signal between the onset of the item ($onset_i$) and its offset($offset_i$):

$$M_{ij} = \frac{\sum_{t=onset_i}^{offset_i} h(t)R_j(t)}{\sum_{t=onset_i}^{offset_i} h(t)} \qquad \text{Eq. (4)}$$

Memory retrieval

At retrieval the context signal is replayed, and each item is re-activated according to the similarity s_i (based on Euclidean distance, d_i) between the current state of R and the stored association M_i:

$$d_i(t) = \sqrt{\sum_j (\mathrm{real}(R_j(t)) - \mathrm{real}(M_{ij}))^2 + (\mathrm{imag}(R_j(t)) - \mathrm{imag}(M_{ij}))^2} \quad \text{Eq. (5)}$$

$$s_i(t) = \min(d) - d_i(t) \quad \text{Eq. (6)}$$

where min(d) refers to the minimum Euclidean distance over all items.

Times corresponding to peaks in similarity are identified, and at each peak the current winner (most active item) is selected for serial output. Activation of each item o_i is calculated taking into account the suppression of previously selected items (q_i) and noise, where ν is a parameter governing noise variance, t is the current time and u_i is the time when the item was previously selected (q_i=0 if the item has not been previously selected), δ is a parameter governing the exponential decay of suppression after selection.

$$\eta \sim \mathcal{N}(0, \nu) \quad \text{Eq. (7)}$$

$$q_i(t, u_i) = 0.5^{(t-u_i)/\delta} \quad \text{Eq. (8)}$$

$$o_i(t) = s_i(t) - q_i(t, u_i) + \eta \quad \text{Eq. (9)}$$

The selected item at each peak is recorded and compared with the target item to generate serial position curves, transposition distances, and item accuracy/error scores for comparison with empirical data. Retrieval of each list structure is simulated 100000 times with the mean proportion of items produced at each serial position calculated for each of the 28 grouping patterns.

Parameters used: filter depth n=15; filter spacing λ=1.2; base frequency f_b=0.1 Hz; filter width σ_b=($1/f_b \times 0.5$)=5 s.

Figure Captions

Figure 1. a) Schematic illustration of an individual amplitude modulation (AM) tuned “neuron” in the BUMP model. The output activity (firing rate) is determined by convolving the neuron’s characteristic impulse response with a signal representing the changing amplitude envelope of an auditory input (see modelling appendix for further details). In this example the neuron’s impulse response has a wavelength of 0.5s corresponding to a frequency of 2 Hz. b) Relationship between input and output for regular and varying rates of AM; the phase of output (bottom axes) tracks the phase of the input (top axes) for both regular (left) and varying rate (right) inputs. The amplitude of the firing rate changes depends on the match between the impulse response tuning and the current rate of amplitude modulation. In this illustration the varying rate signal moves smoothly from 0.9 to 4.0 Hz, with the largest responses being seen as the rate approaches 2Hz.

Figure 2. a) Schematic illustration of a pair of AM tuned “neurons” in the BUMP model. The phases of their impulse responses are shifted by 90 degrees (constituting a complex quadrature filter). b) the phase-offset response of the pair (lower trace) during presentation of a sequence of 9 items represented by triangular AM pulses as in the simulations described in the present paper (upper trace). c) At any time the combined output of the two cells can be represented as a point in 2D space defined by the relative activity of the two “neurons” (gray circle). As time passes the point rotates around the origin. The overall phase of the output can be represented by a hue defined by the angle made with the origin, while its amplitude can be represented by the brightness of the chosen color. d) The evolution of the phase and amplitude of the output over time is illustrated with a colored ribbon. In this instance the cells are tuned to respond to AM modulations with a wavelength of 1.94s which in this example happens to be close to the grouping frequency of the items in the list (b). The phase of the pair’s output goes through approximately one cycle per group.

Figure 3. Phase and amplitude responses of a population of oscillators with different tunings (spaced between 0.1 Hz and 1.28 Hz as in the simulations reported in this paper, see text and modelling appendix for further details). For each axis, the upper trace shows the timing of triangular amplitude pulses representing each item in a nine-item list. The colored phase amplitude diagram represents the evolution of amplitude and phase of oscillators with different tunings (y-axis) over time (x axis). Phase is indicated by hue, amplitude by brightness. a) ungrouped list b) list grouped into three groups of three (3-3-3) c) 2-6-1 grouped list. Note that the same population of oscillators responds to all three sample sequences, but that the phase and amplitude of the entrained responses is systematically affected by the list structure and in particular by local amplitude modulations on different scales (e.g., corresponding to list, group, item). Each item will be associated with the state of the oscillator population at the time it is presented. At retrieval items associated with similar states may be confused with one another, resulting in transposition errors.

Figure 4. Serial position curves from Experiment 1 as a function of list-type, for (a) the no suppression condition and (b) the suppression condition.

Figure 5. Transposition gradients from Experiment 1 as a function of list-type, for (a) the no suppression condition and (b) the suppression condition.

Figure 6. Serial position curves from Experiment 2, averaged across grouping patterns, as a function of list-type.

Figure 7. Scatter plots of the proportion of correct responses for the 28 different patterns of grouping used in Experiment 2 and in Ryan (1969a). Panel (a) plots proportion correct for the unpredictable and predictable conditions of Experiment 2. The remaining panels plot proportion correct for Ryan's

experiment against (b) the predictable condition of Experiment 2 and (c) the unpredictable condition of Experiment 2.

Figure 8. Serial position curves from Experiment 2 as a function of list-type, for four of the 28 patterns of grouping tested: (a) 2-6-1, (b) 4-4-1, (c) 2-3-4, and (d) 3-3-3.

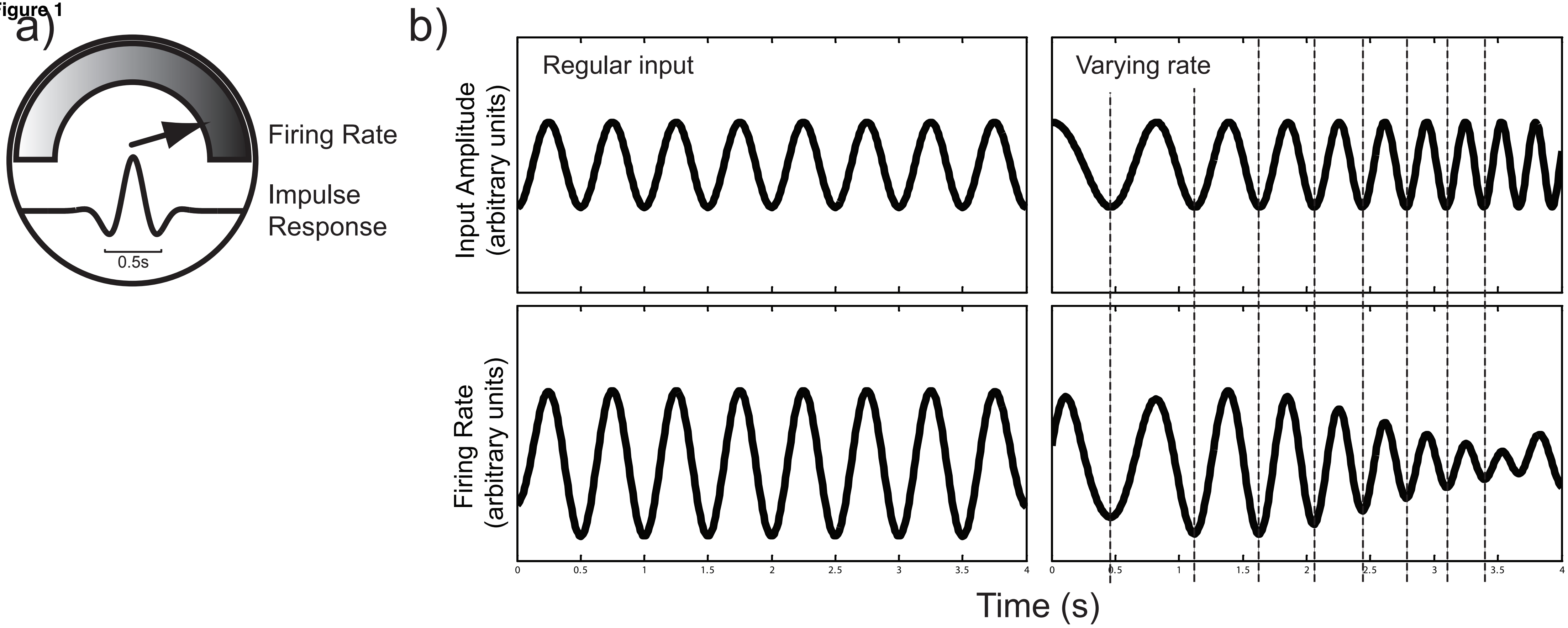
Figure 9. Scatter plots of model simulations of the proportion of correct responses for 28 different patterns of grouping against the observed data of (a) Ryan (1969a), and (b) Experiment 2. See text for further details.

Figure 10. Simulated serial position curves for correct responses and three categories of transposition error alongside observed data from Experiment 2 for four of the 28 grouping patterns (a) 2-6-1, (b) 4-4-1, (c) 2-3-4, and (d) 3-3-3. See text for further details.

Footnotes

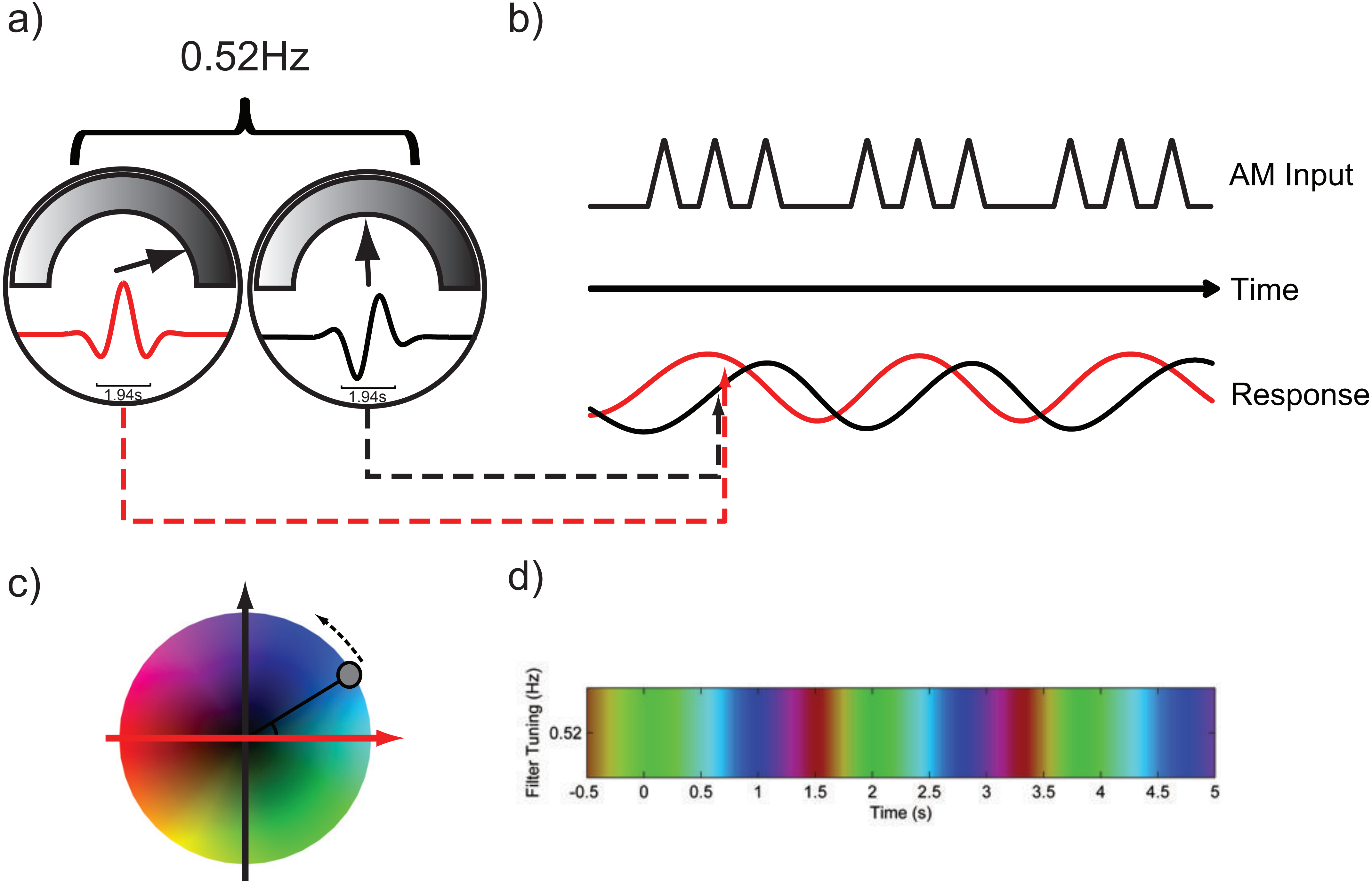
1. This point has been disputed by Botvinick and Plaut (2006), who present a chaining model that reproduces the above effects. However, as we have noted elsewhere (Hurlstone, Hitch & Baddeley, 2014), their model depends on substantial pretraining in which it learns to rely on additional information besides chaining.

Figure 1



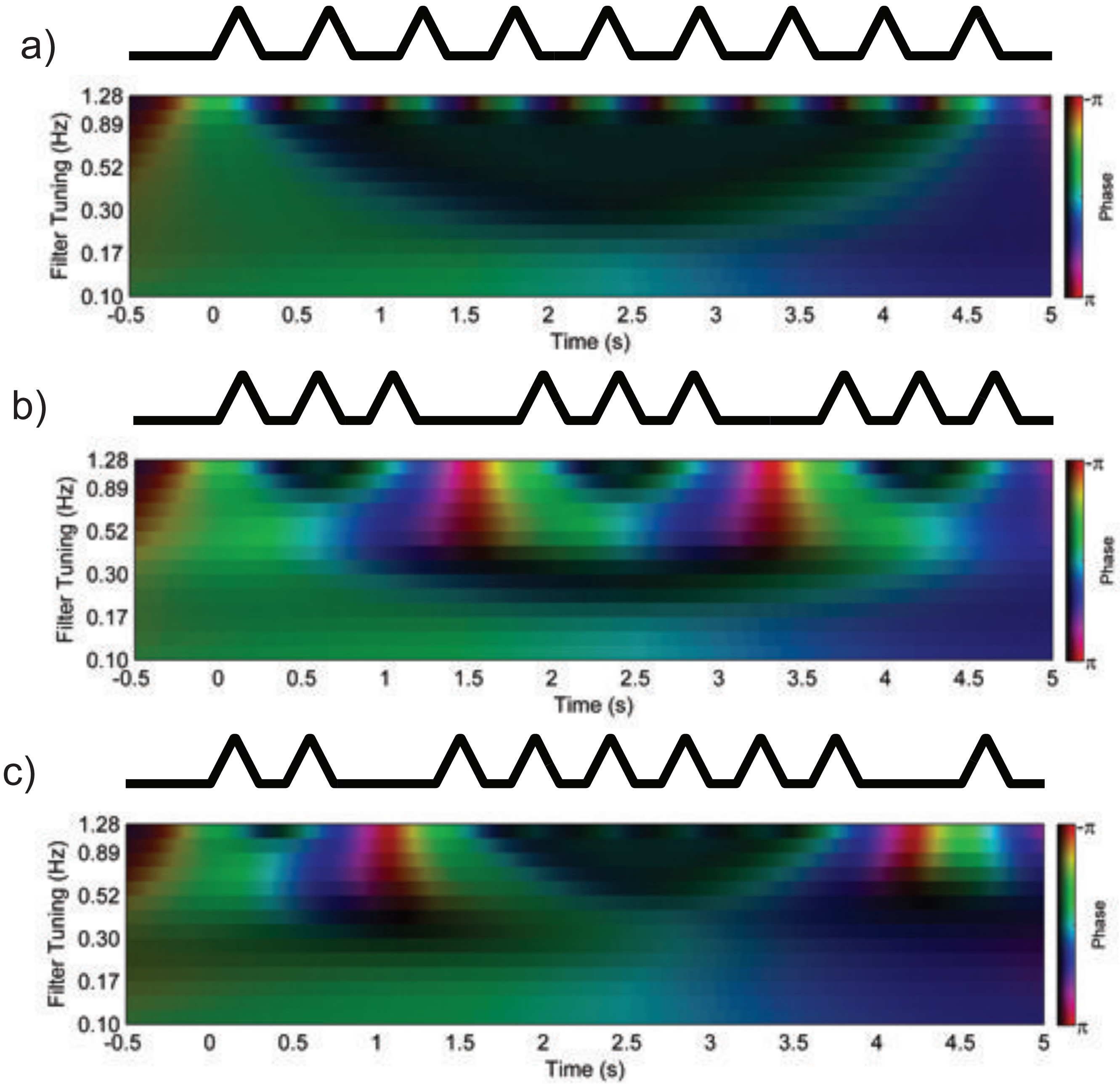
a)
Firing Rate
Impulse
Response
0.5s
b)
Regular input
Varying rate
Input Amplitude
(arbitrary units)
Firing Rate
(arbitrary units)
0
0.5
1
1.5
2
2.5
3
3.5
4
Time (s)

Figure 2



a)
0.52Hz
1.94s
1.94s
b)
AM Input
Time
Response
c)
d)
Filter Tuning (Hz)
0.52
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time (s)

Figure 3



a)
b)
c)
1.28
0.89
0.52
0.30
0.17
0.10
Filter Tuning (Hz)
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
Time (s)
-π
Phase
π

Figure 4

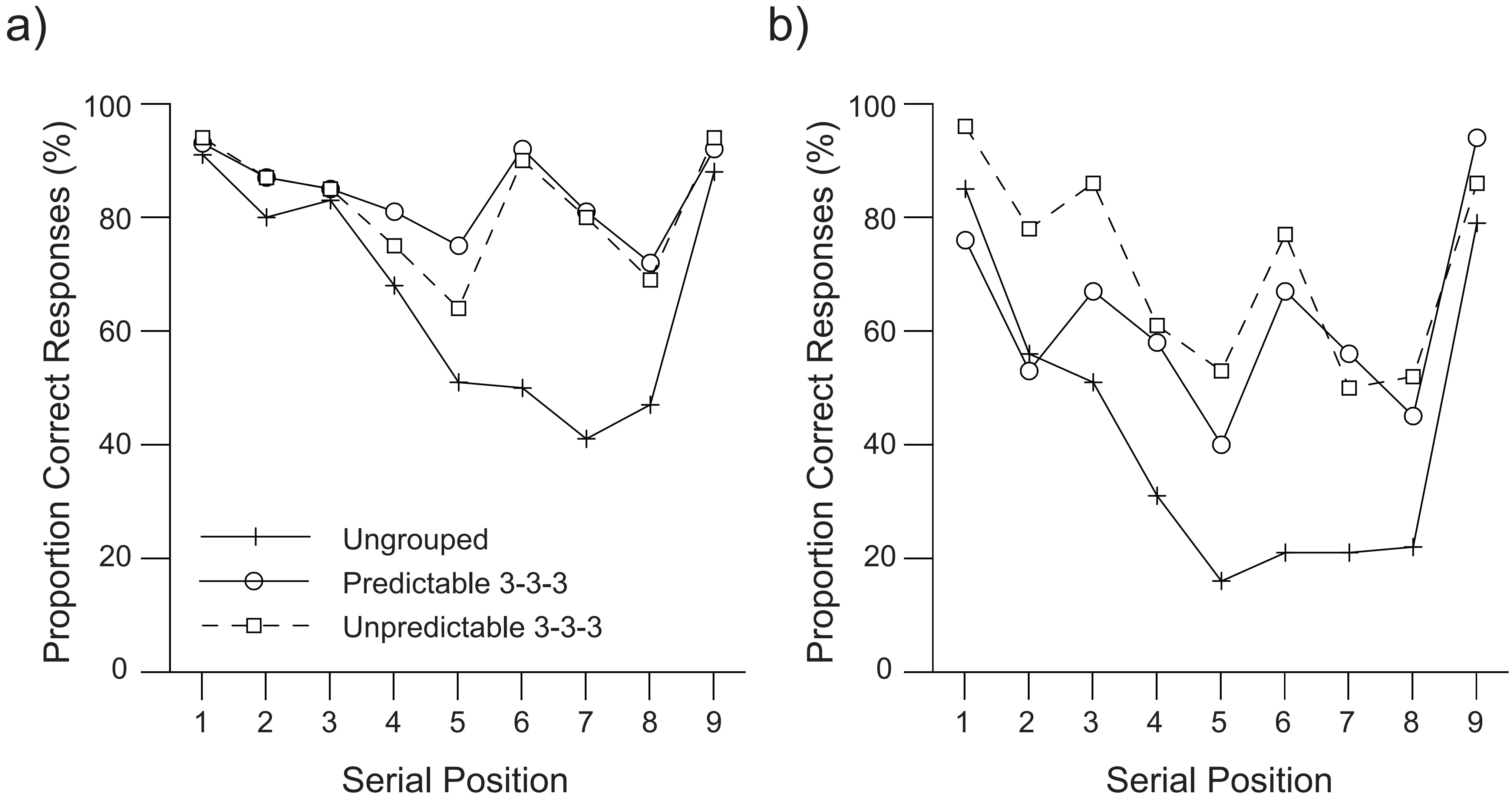
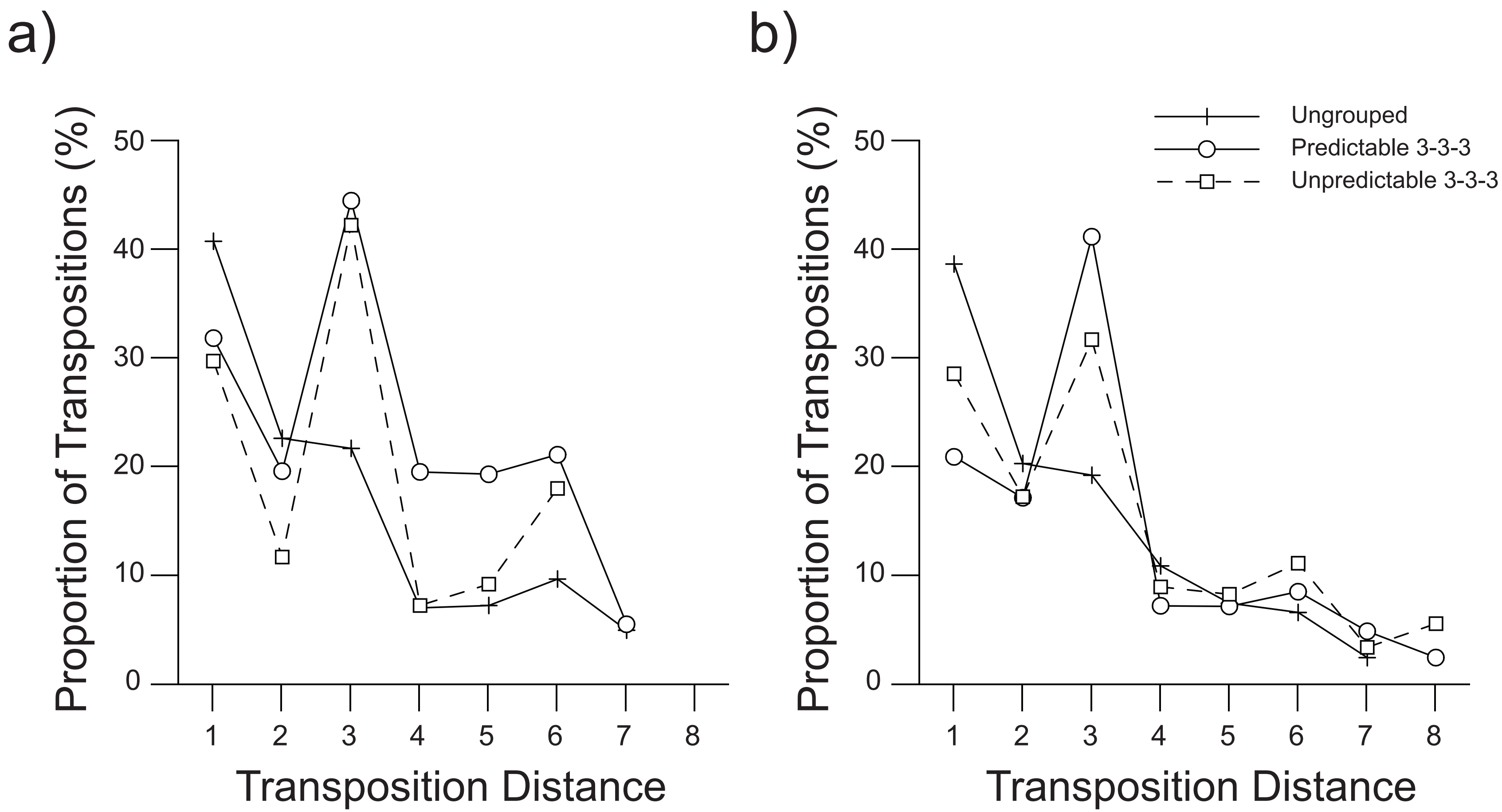
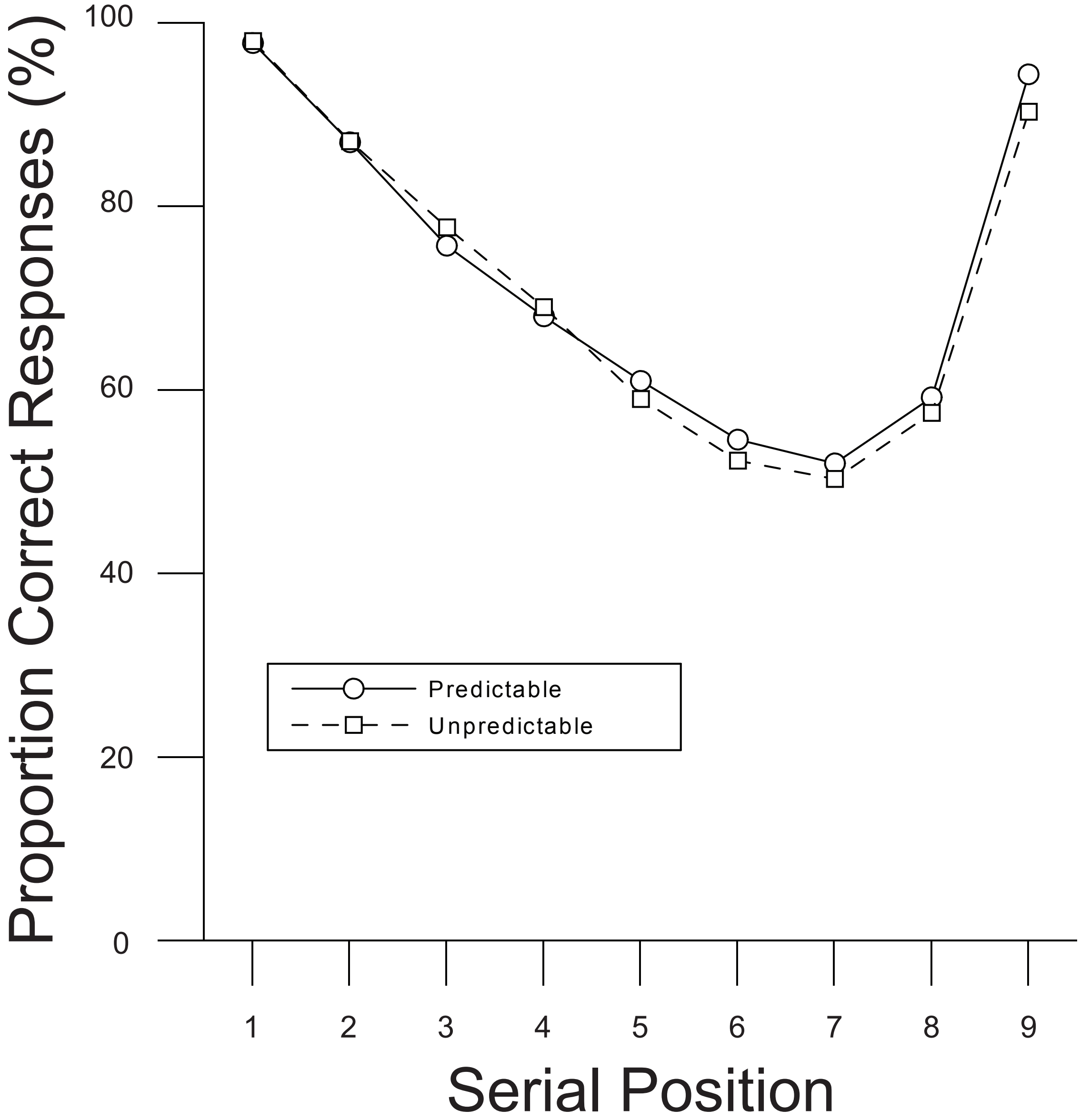

a)
b)
Proportion Correct Responses (%)
Serial Position
Ungrouped
Predictable 3-3-3
Unpredictable 3-3-3

Figure 5



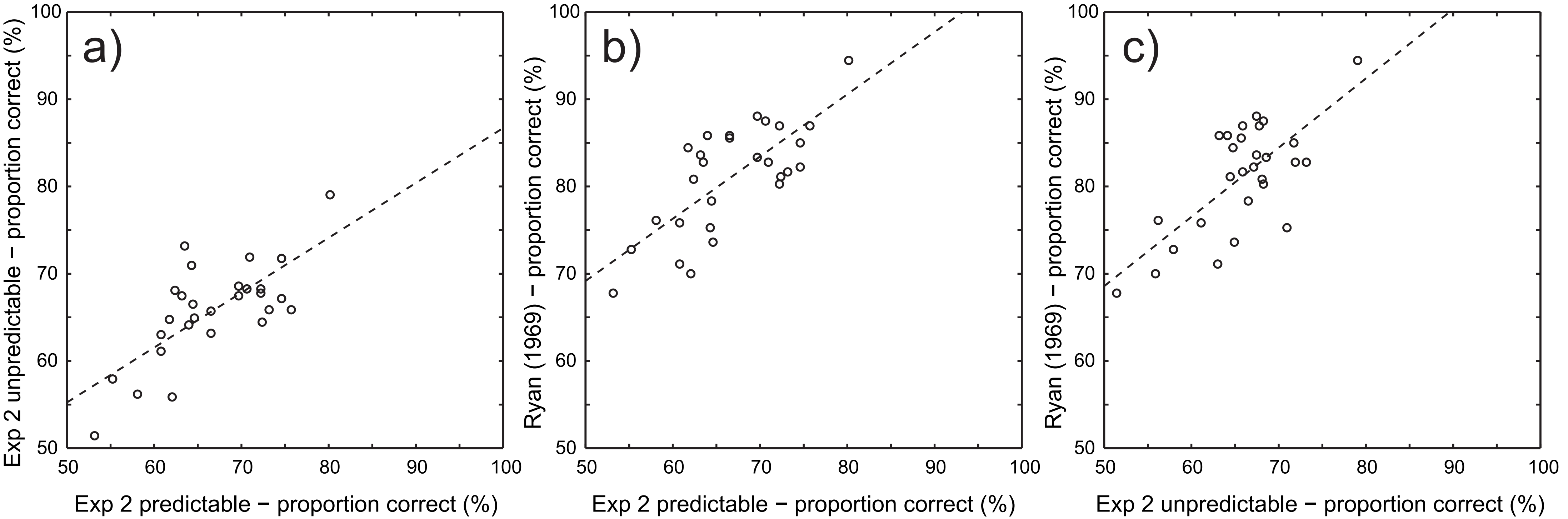
a)
b)
Proportion of Transpositions (%)
50
40
30
20
10
0
1
2
3
4
5
6
7
8
Transposition Distance
Ungrouped
Predictable 3-3-3
Unpredictable 3-3-3

Figure 6



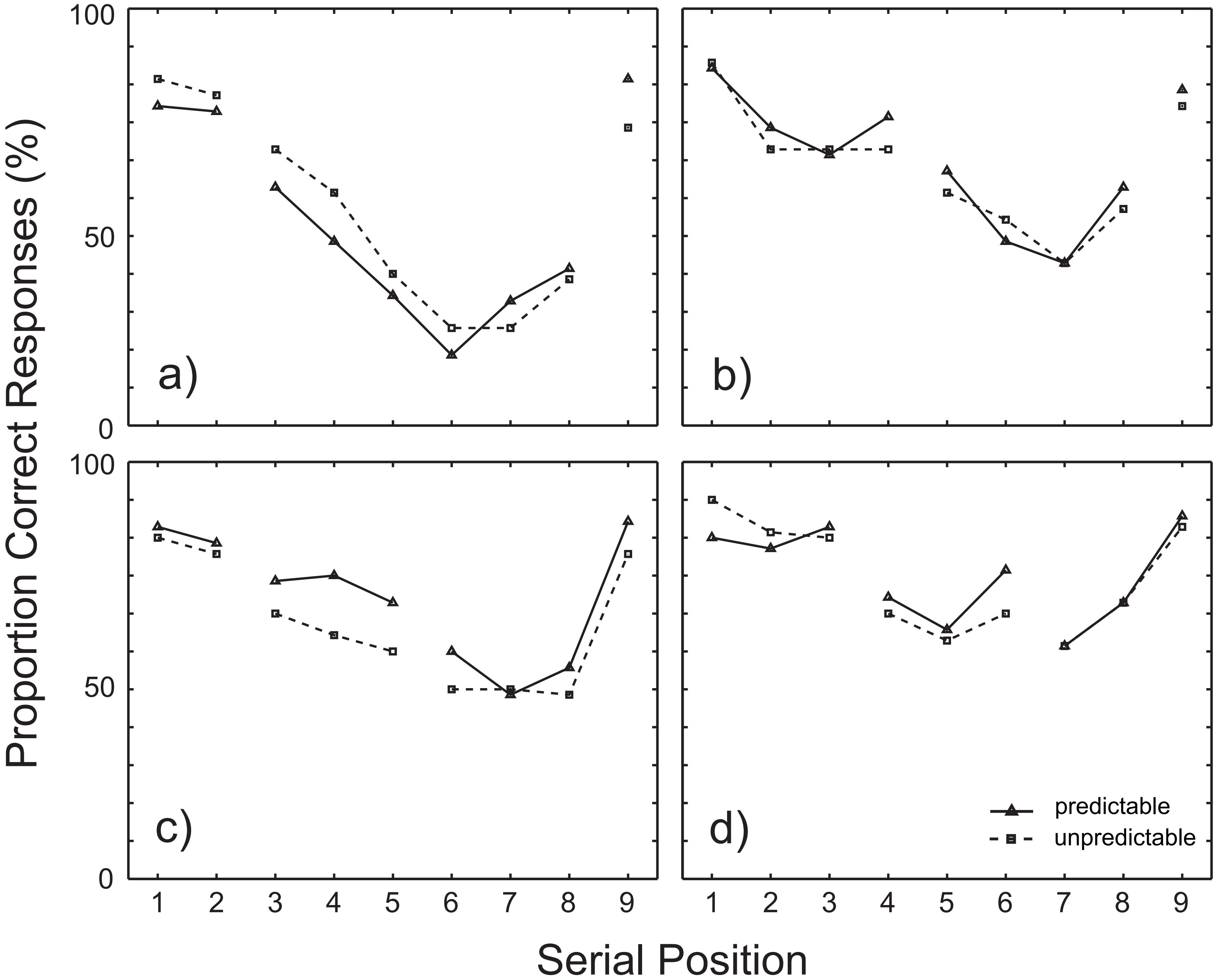
Proportion Correct Responses (%)
100
80
60
40
20
0
Predictable
Unpredictable
1
2
3
4
5
6
7
8
9
Serial Position

Figure 7



a)
b)
c)
Exp 2 unpredictable − proportion correct (%)
Exp 2 predictable − proportion correct (%)
Ryan (1969) − proportion correct (%)
Exp 2 predictable − proportion correct (%)
Ryan (1969) − proportion correct (%)
Exp 2 unpredictable − proportion correct (%)
50
60
70
80
90
100

Figure 8



Proportion Correct Responses (%)
100
50
0
a)
b)
c)
d)
predictable
unpredictable
1
2
3
4
5
6
7
8
9
Serial Position

Figure 9

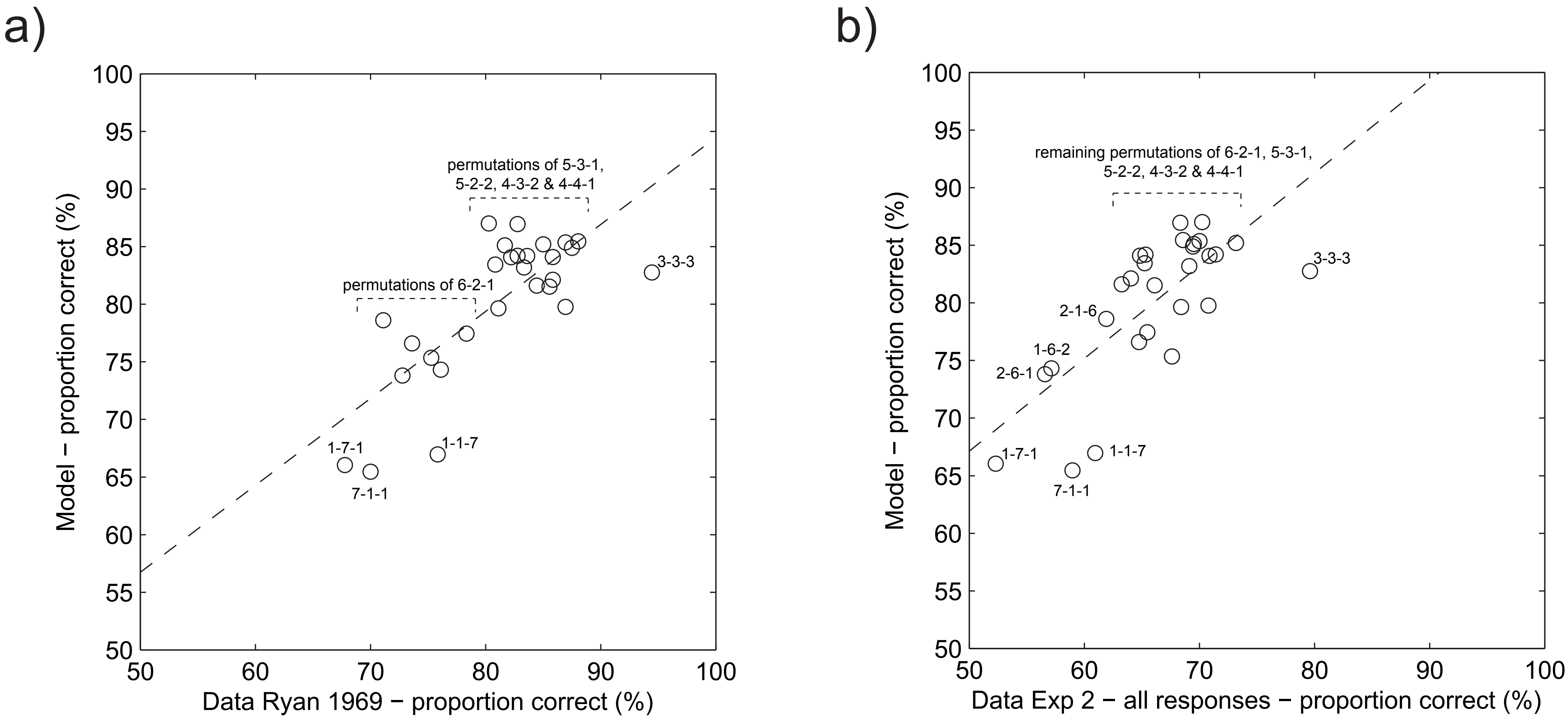
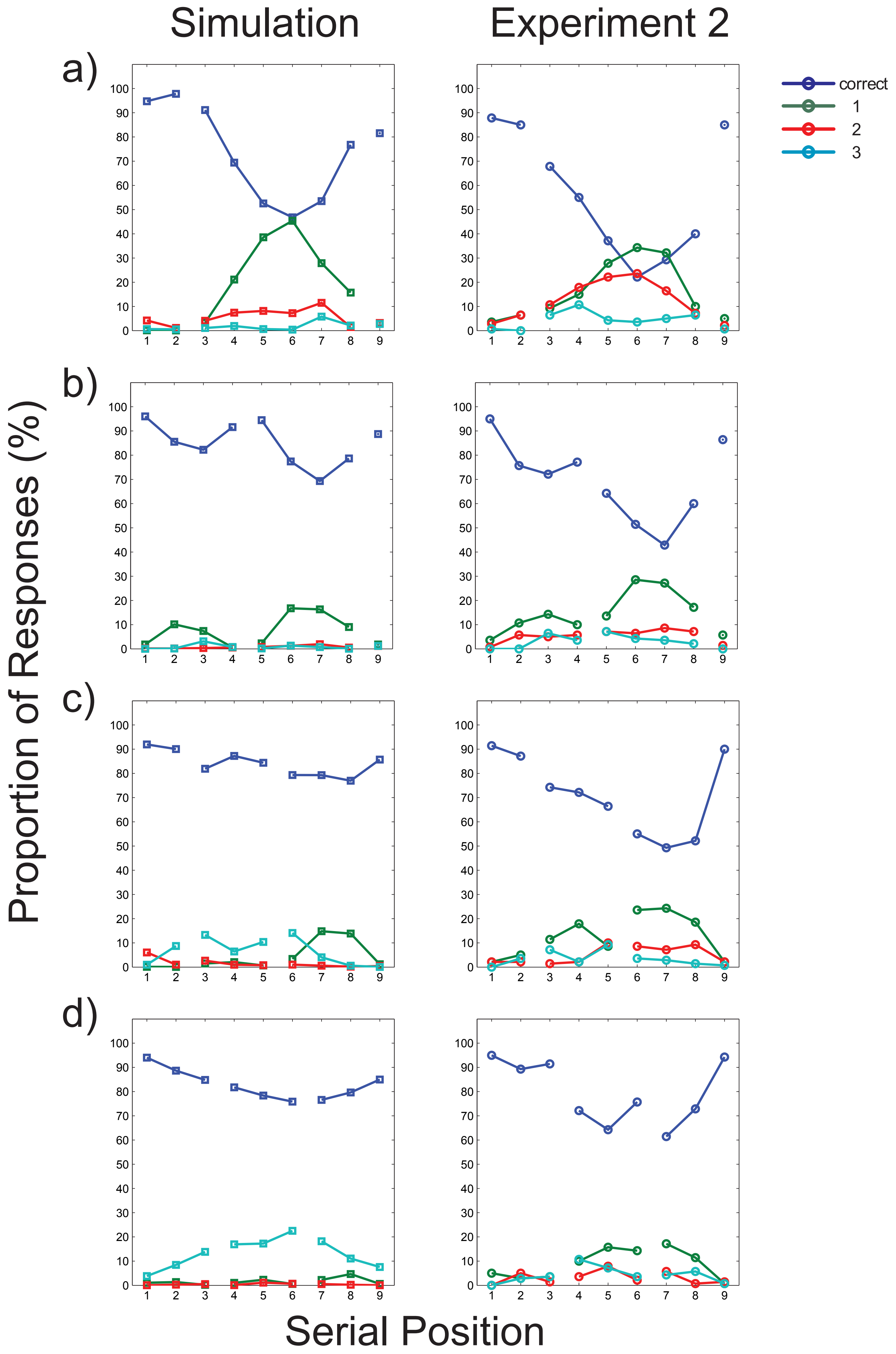

a)
permutations of 5-3-1,
5-2-2, 4-3-2 & 4-4-1
permutations of 6-2-1
3-3-3
1-7-1
7-1-1
1-1-7
Model − proportion correct (%)
Data Ryan 1969 − proportion correct (%)
b)
remaining permutations of 6-2-1, 5-3-1,
5-2-2, 4-3-2 & 4-4-1
3-3-3
2-1-6
1-6-2
2-6-1
1-7-1
7-1-1
1-1-7
Model − proportion correct (%)
Data Exp 2 − all responses − proportion correct (%)

Figure 10



Simulation
Experiment 2
a)
b)
c)
d)
correct
1
2
3
Proportion of Responses (%)
Serial Position

Table(s)

Table 1

Proportions of transpositions within and between groups for Experiment 1 as a function of list-type and secondary-task.

	No Suppression			Suppression		
		Between Groups			Between Groups	
List-Type	Within Groups	Interpositions	Other	Within Groups	Interpositions	Other
Ungrouped	.32	.27	.41	.30	.23	.47
	(.12)	(.18)	(.17)	(.12)	(.11)	(.07)
Predictable 3-3-3	.15	.55	.30	.22	.49	.30
	(.20)	(.31)	(.28)	(.06)	(.14)	(.13)
Unpredictable 3-3-3	.25	.60	.15	.31	.42	.28
	(.11)	(.21)	(.13)	(.11)	(.10)	(.10)

Table 2

Proportion of correct responses in Experiment 2 as a function of list-type and pattern of grouping.

	List-Type			List-Type	
Pattern	Predictably Grouped	Unpredictably Grouped	Pattern	Predictably Grouped	Unpredictably Grouped
117	.608	.611	324	.722	.683
126	.646	.649	333	.802	.790
135	.624	.681	342	.746	.717
144	.722	.678	351	.617	.648
153	.640	.641	414	.732	.659
162	.581	.562	423	.635	.732
171	.532	.514	432	.710	.719
216	.608	.630	441	.706	.683
225	.724	.644	513	.665	.632
234	.746	.671	522	.757	.659
243	.697	.675	531	.697	.686
252	.665	.657	612	.644	.665
261	.552	.579	621	.643	.710
315	.632	.675	711	.621	.559

Supplementary Material

Click here to download Supplementary Material: BUMPSupplement.pdf