Near-Independent capacities and highly constrained output orders in the simultaneous free recall of auditory-verbal and visuo-spatial stimuli.

Cathleen Cortis Mack, Kevin Dent, & Geoff Ward

University of Essex

Correspondence to:

Cathleen Cortis Mack or Geoff Ward,

Department of Psychology,

University of Essex,

Wivenhoe Park,

Colchester, Essex

CO4 3SQ, UK

Telephone: +44 1206 873799 or +44 1206 873800

Fax: +44 1206 873590

E-mail: ccorti@essex.ac.uk or gdward@essex.ac.uk

Running header: immediate free recall of cross-domain stimuli

Abstract

Three experiments examined the immediate free recall (IFR) of auditory-verbal and visuo-spatial materials from single-modality and dual-modality lists. In Experiment 1, we presented participants with between 1 and 16 spoken words, with between 1 and 16 visuo-spatial dot locations, or with between 1 and 16 words *and* dots with synchronized onsets. We found that for dual-modality lists (1) overall performance, initial recalls, and serial position curves were largely determined by the within-modality list lengths, (2) there was only a small degree of dual-task trade-off (which was limited to the visuo-spatial items), and (3) there were strongly constrained output orders: participants tended to alternate between words and dots from equivalent or neighboring serial positions. In Experiments 2 and 3, we compared lists of 6 single-modality items with dual-modality lists of 6 words and 6 dots with synchronous or alternating onsets (Experiment 2), or random but asynchronous onsets (Experiment 3). In all three dual-modality conditions, we again found only a small trade-off in visuo-spatial (but not verbal) IFR performance. There were similarly highly constrained output orders with the synchronous and alternating onsets, and these patterns were present but attenuated with the randomized onsets. We propose that both auditory-verbal and visuo-spatial list items are associated with a common temporal episodic context that is used to guide cross-modal retrieval, and we speculate that the limited, asymmetric interference could arise because the less variable representations of the dots share only a relatively small subset of features with the more variable representations of the words.

249 words

Keywords: free recall; visuo-spatial memory; verbal memory; output order; dual-task interference

This research examined the immediate free recall (IFR) of auditory-verbal and visuo-spatial materials when presented in single-modality and dual-modality lists. There were three main aims of the research: (1) to determine whether verbal and spatial immediate memory share a common capacity limit or exhibit independent or near-independent performance limits, (2) to determine the extent to which the output orders from verbal and spatial memory were independent, and (3) to consolidate and extend our recent work examining the similarities and differences between the output orders and serial position curves in verbal and non-verbal IFR (Cortis, Dent, Kennett, & Ward, 2015).

In a typical IFR experiment, participants are presented with a series of items, one at a time, and at the end of the list, they must try to recall as many items as possible, in whatever order they wish. Historically, the IFR task has tended to use longer lists of between 10 to 40 words as stimuli (e.g., Murdock, 1962; Roberts, 1972), and the list length effect refers to the findings that as the length of the list is increased, so the total number of words recalled increases and the proportion of words recalled decreases (e.g., Murdock, 1962; Roberts, 1972; Ward, 2002). At longer list lengths, participants tend to initiate recall with one of the last few list items (e.g., Hogan, 1975; Howard & Kahana, 1999) and show a tendency for forward-ordered transitions in recall (e.g., Howard & Kahana, 1999; Kahana, 1996). The overall pattern of recall across the different words in the experimenter's list can be illustrated in the bowed serial position curves (e.g., Glanzer & Cunitz, 1966; Jahnke, 1965; Murdock, 1962; Postman & Phillips, 1965) showing primacy effects (enhanced recall of the early list items) and recency effects (enhanced recall of later list items).

Recently, Ward, Tan & Grenfell-Essam (2010) systematically examined the effect of list length on the output order and serial position curves in IFR using shorter lists. They varied the list length unpredictably so that participants saw between 1 and 15 words, but did not know in advance when the list would end. At longer list lengths, participants tended to initiate recall with one of the last few items, and on these trials, the serial position curves were dominated by recency with only modest primacy. By contrast, at much shorter list lengths, participants tended to initiate recall with the first list item, and on these trials, there was greatly reduced recency and elevated recall of the early list items. Ward *et al.* (2010) argued that these findings were important for two reasons: (1) the finding that participants initiate recall of short lists with the first item is difficult to explain by many unitary accounts of IFR (e.g., Brown, Neath, & Chater, 2007; Glenberg & Swanson, 1986; Howard & Kahana, 2002a; Polyn, Norman & Kahana, 2009; Sederberg, Howard & Kahana, 2008; Tan & Ward, 2000) that assume that the end of list items are the most highly accessible at test; and (2) the finding suggests that the natural output order in IFR of short lists closely resembles recall in immediate serial recall (ISR), increasing the prospects of theoretical integration between the two tasks. These findings have been successfully replicated when the list length is both known and not known in advance (Grenfell-Essam & Ward, 2012), and with grouped and ungrouped lists (Spurgeon, Ward, Matthews & Farrell, 2015).

A number of subsequent studies have sought to understand why participants initiate IFR of short lists with the first list item. A rehearsal-based explanation appears unlikely, because the tendency can be obtained under concurrent articulatory suppression and fast presentation rates (Grenfell-Essam, Ward & Tan, 2013). A classic short-term buffer store (STS, Atkinson & Shiffrin, 1971) explanation appears not to be the entire answer either, because the tendency is still present (albeit reduced) under delayed and continual distractor free recall (Spurgeon, Ward & Matthews, 2014b), conditions which might be expected to severely disrupt or eliminate the contents of STS. The finding also appears not to reflect the direct output of a putative phonological store (e.g., Baddeley, 1986) because it is present (albeit reduced) with visual presentation and concurrent articulatory suppression (Spurgeon, Ward & Matthews, 2014a), conditions in which visual verbal items should not be phonologically recoded (Baddeley, 1986). The tendency also survives a stimulus prefix (Grenfell-Essam & Ward, 2015) and can also be observed using visuo-spatial dots and tactile presentations to the face as stimuli (Cortis et al., 2015).

The Cortis et al. findings are particularly important because they ruled out an exclusively language-based explanation of this tendency. Cortis et al. found that participants initiated recall with the first list item at short lists and this tendency decreased with increasing list length for lists of visually presented words, arrays of visuo-spatial rounded rectangles, tactile facial locations, and sequentially presented visuo-spatial dot locations. In addition, the IFR of non-verbal stimuli was not mediated by phonological recoding, because there was no effect of concurrent articulation on non-verbal IFR performance or recall order. Moreover, the Cortis et al. findings provide the most comprehensive set of analyses of the serial position curves and output orders in IFR of visuo-spatial stimuli over a wide range of list lengths.

Previously, only a few studies had examined IFR of nonverbal items. Dent and Smyth (2006) examined overall pattern accuracy in the IFR of 3, 6, 8 and 10 sequential locations. They found a clear effect of list length: the mean accuracy of reproducing the dots in the correct locations decreased with increasing list length and decreased with successive responses. Bonanni, Pasqualetti, Caltagirone, and Carlesimo (2007) also examined IFR sequences of 6, 8, and 10 spatial locations, but these were exemplars from a 25-item grid. Like Dent and Smyth (2006), Bonnani et al. also observed clear list length effects, but they also reported serial position curves. They found primacy effects at all three list lengths, with recency effects developing at the longest list length. They also found that the primacy increased and recency decreased with increasing presentation rates.

More recently, Gmeindl, Walsh and Courtney (2011) examined IFR and ISR of visually-presented digits and spatial locations and used a staircase procedure, starting the list length with only a few items and then increasing (and then decreasing) the list length in line with successful (or unsuccessful) recall performance. They found that ISR performance was superior for digits than spatial locations, but under IFR instructions, digit recall slightly decreased whilst spatial recall greatly increased. Moreover, they found that participants spontaneously recalled both short lists of digits and spatial locations in serial order in the IFR task, and this tendency was twice as likely with the digits than with the locations. With the spatial stimuli, participants sought to reorganize their responses to reduce the spatial distance between clicks, outputting successive responses that were in closer proximity than had been present in the target sequence, perhaps grouping the locations using local spatial configurations (Bor, Duncan, Wiseman, & Owen, 2003). Gmeindl et al. argued that serial order information might be more readily bound to verbal stimuli than spatial stimuli, a finding supported by the greater accuracy in a verbal serial recognition task compared to a spatial serial recognition task.

Finally, Gibson and colleagues in a number of recent papers examined verbal and spatial IFR of 12-item lists in children with ADHD and matched controls (e.g., Gibson, Gondoli, Flies, Dobrzenski, & Unsworth, 2009; Gibson, Gondoli, Johnson & Robison, 2014; Gibson, Gondoli, Johnson, Steeger, Dobrzenski, & Morrisey, 2011). Gibson and colleagues found clear evidence of primacy and recency in both verbal and spatial IFR, and consistent tendency to initiate recall with either the first list item or one of the last 2-3 items. Moreover, Gibson et al. (2009) showed that the resultant serial position curves for both types of stimuli were affected by the first recall: there was greater primacy and reduced recency when recall was initiated with the first item, whereas there was greater recency and reduced primacy when recall was initiated with an end of list item. The Cortis et al. (2015) analyses confirmed the effects of list length, serial position curves, probability of first recall (PFR) curves, and the resultant effect of the first recall on subsequent serial position curves but analyzed recall over a far wider range of list lengths of between 1 and 15 visuo-spatial locations and visual words.

The current three experiments sought to extend the Cortis et al. (2015) findings by making three sets of comparisons between the IFR of single- and dual-modality lists. The first comparison of general interest considers how IFR performance limits that are present in single-modality lists would relate to the IFR performance limits in a dual-modality list. We used auditory-verbal lists of spoken words presented via headphones and visuo-spatial lists of dot locations presented one at a time on a computer screen as stimuli, and at the end of the list, participants recalled the words orally (recorded by a microphone) and recalled the dots by clicking on the screen locations using the computer mouse. If the stimuli from the two modalities were considered to be stimuli from the same (multi-modal) list, then in line with typical studies of list length effects (Murdock, 1962; Roberts, 1972; Ward, 2002), we might expect that participants would recall a slightly increased number of items but with a greatly decreased proportion correct. For example, Roberts (1972) showed that the total number of words recalled didn't greatly increase by doubling the number of words presented at twice the presentation rate (for even less of an effect, see also Waugh, 1967). At the other extreme, if the presentation time in free recall (Roberts, 1972; Waugh, 1967) was still a major determinant of total number of items recalled but it was construed as the proportion of the list *in the same modality* that a given item occupied, then we might find that the IFR of a dual-modality list of say 6 words and 6 dot locations could be additive - equivalent to the independent limits in recalling the two 6-item single-modality lists. This would also be the case if the two classes of stimuli were stored essentially independently – in separate stores or within the same episodic memory with minimal mutual interference.

A second comparison of general interest was in the output orders that participants might adopt when they are free to recall in any order. It is well established that successive outputs in IFR tend to reflect transitions between list items occupying temporally near positions in the experimenter's list compared to more remote positions, and there is a greater tendency to output in forward than in backward recall order (the asymmetric lag recency effect, Howard & Kahana, 1999; Kahana, 1996). Cortis et al. (2015) confirmed that this lag recency effect was present with both verbal lists and sequences of visual-spatial dot locations. Of interest was whether these transitions would similarly be observed for both within-modality and cross-modality list items. If the dual-modality lists were encoded as multimodal integrated lists then one might expect highly constrained output orders between modalities: recall might be expected to initiate from similar contemporaneously-presented list items, and recall might proceed with tightly-coupled outputs from the two modalities. In contrast, previous research has shown that participants might rather prefer to cluster their outputs by stimulus modality. For example, classic early work examining the dichotic presentation for free recall of two separate 3-digit sequences presented to each ear (Broadbent, 1954; for related bisensory serial recall data, see Dornbush, 1968; Margrain, 1967) showed a preferred tendency to recall the items presented to one ear before outputting items presented to the other ear, particularly at faster presentation rates. In addition, when auditory and visual words are presented mixed in a single list, there tends to be an auditory advantage, and clustering by modality during recall (Murdock & Walker, 1969). Moreover, even when lists of words are presented for IFR from different semantic categories, there is clear evidence of semantic clustering at recall (e.g., Bousfield, 1953) – indeed, there are semantic contiguity effects even from nominally unrelated list items (Howard & Kahana, 2002b). It is a reasonable alternative hypothesis that participants might cluster their recall by modality, and might even initiate recall of stimuli from different modalities with items that were encoded in very different serial positions. This would also be the case if the two classes of stimuli were stored essentially independently: the output order from one store, source or mechanism may be independent from the other.

Finally, a third set of comparisons further examined the similarities and differences between the IFR of auditory-verbal word lists and visuo-spatial dots. Across all three experiments, it will be possible to compare the serial position curves and output orders in the two single-modality conditions. In addition, Experiment 1 allowed us the possibility to extend Cortis et al. (2015) by examining specifically whether the effect of list length on the first recall from a dual-modality list was similar to the effect of list length on the first recall from the corresponding single-modality list. Following from the work of Ward et al. (2010) and Cortis et al. (2015), we continue to be interested in why participants change their modal tendency to initiate recall from the first list item at shorter lists to initiate recall of the one of the most recent items with longer lists. If the change or cross-over in the item first recalled with increasing list length is driven by a performance limit in holding a particular number of items in mind, then one might see the tendency to initiate recall with the first list item decrease dramatically with increasing list length in one or both modalities with dual-modality lists relative to single-modality lists. By contrast, if the initial tendency is not driven by a capacity limit or if the capacity limits of verbal and non-verbal items are essentially independent, then there should be little or no effect on initial outputs when participants are required to output from two rather than just a single modality.

Unfortunately, to date, it is difficult to predict from the existing IFR literature whether verbal and non-verbal items might be expected to show a dual-task decrement in performance. Most, if not all, theorizing in IFR has been based on explaining the recall of word lists (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998), and it is therefore not surprising that many accounts of IFR feature language-specific mechanisms and seek to explain language-specific effects. For example, many contemporary accounts: (1) include long-term lexical and semantic representations (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann & Usher, 2005), (2) seek to explain semantic clustering effects (e.g., Healey & Kahana, 2014; Howard & Kahana, 2002b; Polyn et al., 2009), and (3) assume verbal control processes such as rehearsal (e.g., Laming, 2006, 2008, 2010; Raaijmakers & Shiffrin, 1981; Tan & Ward, 2000). Although some accounts of IFR (e.g., Atkinson & Shiffrin, 1971) assumed that there are separate modality-specific sensory stores (e.g., Crowder & Morton, 1969; Sperling, 1960) and detailed the role of recoding and other control processes in STS (e.g., Atkinson & Shiffrin, 1971), the majority of IFR models assume that storage in STS or primary memory (PM, Unsworth & Engle, 2007) and long-term memory is either verbal or amodal. Although most theories would predict that the degree of interference between verbal and visuo-spatial items should be less than the degree of interference between same modality items (e.g., Brown, et al., 2007; Cowan, 1988, 1998, 2001), there are few, if any, contemporary theories of IFR that have explicitly assumed that verbal and visuo-spatial memory items are represented in very different memory stores or systems (but for an exception, see Paivio, 1971, 1991), nor have many contemporary theories of IFR been directly applied to non-verbal IFR data.

Perhaps the most relevant prior empirical research comes from the serial recall literature, which has often considered the effects of trying to maintain two sets of verbal or sets of visual and verbal memory items concurrently. For example, Sanders and Schroots (1969) presented participants with a set of 6 consonants (or digits or tones) to recall in serial order that was immediately followed by a second set of between 2 and 6 visual consonants, 2 and 6 visual digits, between 2 and 6 tones, or between 1 and 5 spatial locations that was also to be remembered and recalled. Sanders and Schroots (1969) found that serial recall improved when the successive sublists came from distinct categories. Although there was a relatively small effect of a change in category (e.g., between visual consonants and digits), there were far greater effects when the list changed from consonants to tones or spatial arrays. More recent work by Morey and colleagues (e.g., Cowan & Morey, 2007; Morey & Mall, 2012; Morey & Miron, 2016; Morey, Morey, van der Reijden & Holweg, 2013; see also Depoorter & Vandierendonck, 2009; Vandierendonck, 2016) suggests that the cross-modal interference involving verbal serial recall may be asymmetric. Visuo-spatial memory appears to be more susceptible to the requirement to maintain a concurrent verbal serial memory load, whereas verbal serial memory appears to be relatively unaffected by a concurrent visuo-spatial memory load.

Within the serial recall and working memory literatures, these dual-task experiments are often conducted within an ongoing debate as to whether or not there are distinct domain-specific memory stores or subsystems (with similar recall mechanisms). There appears to be growing consensus that there is some *functional equivalence* in the ways in which visuo-spatial and verbal stimuli are recalled in ISR (e.g., Guérard & Tremblay, 2008; Hurlstone, Hitch, & Baddeley, 2014; Parmentier, 2011; Ward, Avons & Melling, 2005). However, theorists disagree as to whether this equivalence should be attributable to highly similar but separate modality-specific stores or suggests common domain-general memory processes. For example, the classic Working Memory model and variants (e.g., Baddeley, 1986, 2000, 2012; Baddeley & Hitch, 1974; Baddeley & Logie, 1999; Doherty & Logie, 2016; Logie, 2011) propose that verbal items are represented in a phonological store that is refreshed by verbal rehearsal within the articulatory loop (supported by long-term language knowledge, Baddeley, 2000), whereas visuo-spatial items are represented in the visuo-spatial sketchpad or visual cache, which itself may be refreshed by domain-specific mechanisms (Logie, 1995; Smyth & Scholey, 1994) and supported by long-term visual semantics (Baddeley, 2000). Hurlstone et al. (2014) has recently summarized the case for functional similarities in ISR across domains, but proposed separate verbal and visual memory stores based on (1) studies of dual-task interference showing the selective interference by secondary tasks that are from similar relative to different modalities, (2) the double dissociations observed in neuropsychological patients with complementary patterns of preserved and impaired abilities, and (3) the distinct modality-specific localizations in neuroimaging data. Alternative explanations for the patterns of dual-task interference emphasize differential interference between more similarly- or less similarly-encoded items within domain-general models of memory (Brown, et al., 2007; Cowan, 2005; Nairne, 1990; Neath, 1999; Oberauer, 2009; Oberauer & Kleigl, 2006) or emphasize the extent to which there are overlapping cognitive processes associated with stimulus perception and response production in the two tasks (e.g., Jones, Hughes, & Macken, 2006; Jones, Macken, & Nichols, 2004).

One methodological advantage in our design is that the presentation of auditory-verbal and visuo-spatial lists can occur in parallel: we use spoken words presented via headphones and visuo-spatial dot locations presented on a computer screen. Similarly, recall of the two stimulus types can also occur in parallel: we request oral recall of the words and pointing recall of the visuo-spatial dot locations using a computer mouse. Finally, the two stimulus sets do not obviously share the same representations (e.g., only words require lexical access).

A second advantage of using an IFR methodology over previous serial recall studies, is that we can examine participants' preferred orders of responses when they are attempting to recall all the items that they can from both stimulus modalities, *in whatever order they wish.* Thus, we can see whether participants prefer to treat a dual-modality list as essentially two separate single-modality lists (in which case they may recall by modality in independent output orders) or whether participants may instead prefer to treat a dual-modality list as a single, integrated multi-modal list and respond with successive cross-modal transitions from neighboring temporal or serial positions.

We recognize that it is difficult to make confident predictions from current theories of ISR or working memory to the proposed studies of single-domain and dual-domain IFR, since most of these theories have yet to be applied to the IFR of non-verbal items (Cortis et al., 2015). Moreover, within each of the different classes of memory theories, there are variations or versions that could predict a wide range of different findings. Our research strategy is therefore to highlight a range of potential possibilities, present theoretically interesting data sets before considering how such data constrain current approaches.

Experiment 1

Experiment 1 compared the IFR of single-modality lists of verbal and visuo-spatial material to the IFR of concurrently presented lists of both verbal and visuo-spatial stimuli. There were three groups of participants. Participants in the *words only* group were presented with 50 lists of between 1 and 16 auditory words that were delivered over headphones and participants were required to recall out loud. Participants in the *dots only* group were presented with 50 lists of between 1 and 16 visuo-spatial dot sequences presented silently on a computer screen and participants were required to recall the dot locations by clicking on the locations using a computer mouse. Finally, the *synchronized words and dots* group were presented with 50 lists of dual-modality lists items containing between 1 and 16 auditory words presented with synchronized onsets with the corresponding number of visuo-spatial dot locations. In each group, there were ten trials of each of the following five different within-modality list lengths: 1, 2, 4, 8, and 16, and for each participant the presentation order of the trials with different list lengths was randomized, so that participants did not know in advance when the list was to end.

Experiment 1 addressed all three aims of the research. First, we wished to compare the performance limitations in recalling the auditory-verbal and visuo-spatial stimuli in single and dual-modality lists. Second, we wished to compare the preferred output orders in recalling the single and dual-modality lists. Finally, we wished to extend the findings observed by Cortis *et al.* (2015) who showed highly similar effects of list length on the output order and serial position curves in IFR of verbal and visuo-spatial stimuli. Specifically, we wished to compare how participants tended to initiate recall of the auditory verbal and visuo-spatial lists with increasing list length in the single- and dual-modality lists.

Method

Participants. A total of 60 students from the University of Essex participated in exchange for either course credit or a small payment. The three experiments were undertaken with the approval of the Research Ethics Board of the University of Essex, GW1205 and GW1104.

Materials and Equipment. The auditory-verbal stimuli consisted of the digitized voice files of all the words from the Auditory Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). These .voc formatted files were downloaded from Michael J. Kahana's Computational Memory Lab website (http://memory.psych.upenn.edu/Word_Pools) and converted to .wav files. The words were spoken in a female voice in a US accent, and the files were of the same temporal duration. There were a total of 484 different files, of which a different random 414 were presented to each participant. The visuo-spatial stimuli consisted of circular black dots that were 35mm in diameter. The dots were centered at one of 414 different spatial locations on the screen (arranged in 18 rows by 23 columns), within a 375mm x 280mm frame. The grid was not visible to participants who simply saw 35 mm diameter black dot circles appear on a gray background screen. The stimuli were presented using the application, Supercard, on an Apple Macintosh Computer and the spoken responses were recorded using the application, Audacity.

Design. The experiment used a mixed factorial design. The between-subjects independent variables were the stimulus group with three levels: words alone group, dots alone group, and synchronized words and dots group. There were two within-subjects independent variables: list length, with 5 levels (1, 2, 4, 8, and 16 within-modality items), and serial position with up to 16 within-modality levels. The main dependent variable was the proportion of items correctly recalled, but the distribution of words first recalled and the output orders were also examined.

Procedure. Each participant was randomly allocated to one of the three experimental groups. Participants in the words only group heard spoken words via headphones for oral recall, participants in the dots only group saw a sequence of visuo-spatial dots on a computer screen for recall using a computer mouse, and participants in the synchronized words and dots group received and recalled both sequences of stimuli that were presented with synchronized onsets.

All participants were informed that they would be shown two practice trials of 8 stimuli followed by 50 experimental trials of 1, 2, 4, 8, and 16 stimuli. The 50 experimental trials were arranged into two blocks that each consisted of 5 trials of each of the 5 different list lengths in a random order. Each trial started with a “Ready for next trial?” visual prompt that allowed participants to initiate the list presentation whenever they were ready to do so. Following a computer mouse click, participants were presented with between 1 and 16 stimuli presented one at a time in the single-modality groups, or between 1 and 16 pairs of stimuli for the synchronized words and dots group, where the stimulus onsets of the words and dots were synchronized. The stimuli were presented for 0.75s, with a 0.25s inter-stimulus interval. During the presentation of all types of stimuli, the location of the mouse cursor was locked to a location at the right hand edge of the screen to prevent participants from using it as an external memory aid to a particular location.

At the end of the list there was a tone that prompted recall. The participants in the auditory-verbal (words alone) group said out loud as many words as they could remember, in any order that they wished. Their auditory-verbal responses were recorded to enable later scoring and analyses. The participants in the visuo-spatial (dots alone) group clicked on the locations of the screen where they had previously seen the dot circles, and they were also free to respond in any order that they liked. The maximum number of responses on any trial was equal to the number of presented stimuli, but participants were free to omit responses and many participants made far fewer responses than the maximum. The participants in the synchronized words and dots group recalled both the words and the visuo-spatial locations in the same way as the participants in the single-modality group, and they were instructed to recall as many items from both modalities as possible in any preferred order. After the participants were satisfied that they had completed their recall, they pressed the "submit" button, and this initiated the next trial.

Results

Overall Accuracy. Figure 1 shows the overall accuracy at recalling the two types of stimuli as a function of the list length in the single- and dual-modality lists. A word was scored as correctly recalled only if it was identical to the word that was presented. A dot location was scored as correctly recalled only if the location of the mouse click fell within the circumference of the dot. It was necessary to adopt this rather strict criterion because there were up to 16 possible dot stimuli on a trial, and each 35mm diameter dot occupies approximately 0.9% of the total 375mm x 280mm area. Thus, a random location clicked in a 16-dot sequence would be correct on approximately 15% of occasions.

Considering first the recall of the auditory-verbal items, a 2 (group: words only and synchronized words and dots) x 5 (list length: 1, 2, 4, 8, and 16 words) analysis of variance (ANOVA) was performed on the proportion of words correctly recalled. There was a non-significant main effect of group, F (1, 38) = 0.243, MSE = .009, η^2_p = .006, p = .625, a significant main effect of within-modality list length, F (4, 152) = 1548, MSE = .003, η^2_p = .976, p < .001, and a non-significant interaction, F (4, 152) = .889, MSE = .003, η^2_p = .023, p = .472. Thus, participants showed a clear list length effect when recalling lists of words, and participants' accuracy at recalling the words was essentially unaffected by whether or not they were also tasked to encode and recall the visuo-spatial dots.

--Figure 1 about here--

Considering next the recall of the visuo-spatial items, a 2 (group: dots alone and synchronized words and dots) x 5 (list length: 1, 2, 4, 8, and 16 dots) ANOVA was performed on the proportion of dots correctly recalled. There was a non-significant main effect of group, F (1, 38) = 2.31, MSE = .031, η^2_p = .057, p = .137, a significant main effect of within-modality list length, F (4, 152) = 329.1, MSE = .010, η^2_p = .896, p < .001, and a non-significant interaction, F (4, 152) = 1.00, MSE = .010, η^2_p = .026, p = .408. Thus, participants showed a clear list length effect when recalling lists of dots, and participants' accuracy at recalling the dots using the strict scoring criteria was essentially unaffected by whether or not they were also tasked to encode and recall the words. It should be noted that participants in the dots alone group made significantly more mouse clicks (5093) than those in the synchronized group (3800), t(38) = 4.61, p < .001, d = 1.45, suggesting that at least some of the numeric (but non-significant) advantage for the dots only condition may reflect increased tendency to make additional mouse clicks.

Finally, two comparisons were made confirming what can be seen in Figure 1, that the recall for words was better than the recall for dots. An independent samples t-test showed that the recall of words in the words only group was greater than the recall of the dots in the dots only group, t(38) =

11.36; $p < .001$; $d = 3.59$. Similarly, a paired-samples *t*-test also showed that words were recalled better than dots in the synchronized words and dots group, $t(19) = 8.69$; $p < .001$; $d = 3.22$.

Serial Position Curves. Figure 2 shows the serial position curves for each of the five list lengths. Figure 2A shows the auditory-verbal data from the words alone and synchronized words and dots groups; Figure 2B shows the visuo-spatial data from the dots alone and synchronized words and dots groups.

--Figure 2 about here--

The serial position curves for first the words recalled and then the dots recalled were analyzed at each list length, using a series of 2 (group: single-modality and dual-modality) x *n* (serial positions: 1 to *n*) mixed ANOVAs (where *n* is list length). The exact statistics for the main effects and interaction for each list length can be found in Appendix A. Summarizing first the serial position curves for the recall of the words, the main effects of group on verbal performance were non-significant at all list lengths. There were significant main effects of serial position at list lengths 4, 8 and 16 (list length 4 showed primacy effects, list length 8 showed both primacy and recency effects, whereas list length 16 predominantly showed extended recency effects). Furthermore, the interactions between group and serial position were significant only for list lengths 8 and 16, reflecting a slightly decrease in primacy at these list lengths in the synchronized words and dots group. Summarizing next the serial position curves for the recall of the dots, there were non-significant main effects of group for all but list length 16. There was a significant main effect of serial position for list lengths 8 and 16 (list length 8 showed equal amounts of primacy and recency, while at the longest list length there was extended recency), and there were no significant interactions.

Finally, we compared the shapes of the serial position curves in the two modalities. A set of between-subjects comparisons compared the serial position curves at list lengths 2, 4, 8, and 16 of the words alone group with the corresponding serial position curves of the dots alone group. Recall performance was significantly higher for words than for dots at all list lengths (all *F*s > 12). In addition, there was a significant interaction due to enhanced recency effects with words relative to dots at list length 8, $F(7, 266) = 3.33$, $MSE = .036$, $\eta^2_p = .081$, $p <.002$, and a significant interaction due to enhanced primacy and recency effects with words relative to dots at list length 16, $F(15, 570) = 8.51$, $MSE = .024$, $\eta^2_p = .183$, $p <.001$. An essentially identical set of findings was observed in the within-subjects comparisons between the serial position curves of the words with and the dots at list lengths 2, 4, 8, and 16 in the with synchronized words and dots group.

Probability of First Recall Data. The Probability of First Recall (PFR) refers to the proportion of trials in which the first item recalled was of a particular input serial position (for related analyses, see Hogan, 1975; Howard & Kahana, 1999; Laming, 1999). The full distribution of the items first recalled for each list length, stimulus type, and group are presented in Table 1. Consistent with Ward et al. (2010), the bold values highlight that participants tend to initiate their recalls with the first list item on short lists, but with one of the last four list items with longer lists.

--Table 1 about here--

Figure 3 illustrates how the tendency to initiate recall with the first list item decreases with increasing list length for the auditory-verbal and visuo-spatial data from the single- and dual-modality lists. Consistent with Cortis et al. (2015), participants tended to initiate recall with the first list item at the shorter list lengths but this tendency decreases with increasing list length for both the words and the dots. The decline in recall of the first list item with increasing list length appears steeper for the dots than the words, but a closer look at Table 1 shows that much of the decline is because participants make inaccurate first responses at longer list lengths.

--Figure 3 about here--

Importantly, these findings extended to the novel conditions where participants had to recall both the spoken words and the dots, and the tendency to initiate recall with the first list item was non-significantly reduced in the dual-modality lists. Considering first the auditory-verbal stimuli, a 2 (group: words alone, words from synchronized words and dots) x 5 (list length: 1, 2, 4, 8, and 16 words) ANOVA was performed on the tendency to initiate recall with the word presented in serial position 1. There was a non-significant main effect of group, *F* (1, 38) = 1.98, *MSE* = .071, η^2_p = .049, *p* = .168, there was a significant main effect of within-modality list length, *F* (4, 152) = 214, *MSE* = .026, η^2_p = .849, *p* < .001, and a non-significant interaction, *F* (4, 152) = 1.89, *MSE* = .003, η^2_p = .047, *p* = .115. Similarly, a 2 (group: dots alone, dots from synchronized words and dots) x 5 (list length: 1, 2, 4, 8, and 16 dots) ANOVA was also performed on the tendency to initiate recall with the location of the first dot in the sequence. Again, there was a non-significant main effect of group, *F* (1, 38) = .513, *MSE* = .047, η^2_p = .013, *p* = .478, there was a significant main effect of within-modality list length, *F* (4, 152) = 158, *MSE* = .025, η^2_p = .806, *p* < .001, and a non-significant interaction, *F* (4, 152) = .843, *MSE* = .025, η^2_p = .022, *p* = .500.

Detailed analysis of output order of the words and dots from the synchronized words and dots group. In this section, we consider in more detail the output orders from the synchronized words and dots group, by considering the order of outputs both within- and across- modalities. There are four main findings. First, even though participants were simultaneously presented with equal numbers of words and dots, they nevertheless preferred on over 80% of trials to initiate recall with a word. Specifically, for list lengths 1, 2, 4, 8, and 16, the proportions of trials in which words were first recalled were: .72, .82, .84, .84, .81, respectively (for all list length, *p*s < .001).

Second, on any given dual-modality trial, there was a tight coupling between the first word that was recalled from the auditory-verbal list and the first dot that was recalled from the visuo-spatial list. Table 2 shows the relationship between the first word recalled and the first dot recalled for each of the different list lengths in the dual-modality group. At each list length, it is clear that the within-modality serial position of the first response in one modality often closely matches that of the first response of the other modality. For example, at list length 4, 41.5% of participants' first responses in each modality were the first presented auditory word and visuo-spatial dot respectively, whereas at list length 16, 29% of participants' first responses in each modality were the last presented auditory word and visuo-spatial dot respectively. We calculated the proportion of tightly-coupled first recalls in which the first word recalled from the synchronized words and dots group was from the same serial position as the first dot recalled. For list lengths 2, 4, 8, and 16, the proportions of tightly coupled first responses were never lower than .75: 132/136 (.97), 96/104 (.92), 93/113 (.82) and 86/117 (.75), respectively.

--Table 2 about here--

Third, participants tended to transition between modalities more often than they transitioned within modalities. Figure 4 shows the frequencies of transitions in the synchronized words and dots group. Of the 1,912 recorded transitions from list length 4, 8, and 16, 24% were within modality transitions (344, 18%, Word-to-Word transitions, WW, and 113, 6% Dot-to-Dot transitions, DD), in contrast, 76% were within modality transitions (782, 41% Word-to-Dot (WD) and 673 (35%) Dot-to-Word transitions, DW).

--Figure 4 about here --

To calculate the lag transitions, the within-modality serial position of output *n* + 1 was deducted from the within-modality serial position of output *n*. For example, if a participant recalled the first word in the list (serial position 1), followed by the first dot in the list (serial position 1) this would result in a WD transition of lag 0 (1-1 = 0). Similarly, a transition from the first dot to the second word would be a DW transition of lag +1 (2-1 = +1), whereas a transition from the seventh dot to the fourth dot would be a DD transition of lag -3 (4-7 = -3). As Figure 4 demonstrates, when participants are asked to recall concurrently presented words and dot in any order, they are most likely to recall a word followed by the dot of the same within-modality serial position (WD lag 0), and then transition from the dot to the next word (DW lag +1).

Discussion

Experiment 1 sought to address all three main aims of the research. First, we compared the performance limitations in recalling auditory-verbal and visuo-spatial stimuli in single lists with those in recalling the same material in dual-modality lists. We observed that participants attempting IFR with dual-modality lists of auditory-verbal words and visuo-spatial dots recalled almost as many words and dots as participants who attempted IFR with equivalent single-modality lists. In terms of IFR data, dual-modality performance differed from classic single-modality verbal performance data that showed relatively little increase in the number of words recalled when twice as many words were presented in the same total presentation time (e.g., Roberts, 1972; Waugh, 1967). Instead, our findings suggest that the auditory-verbal and visuo-spatial items in dual-modality lists are maintained with little or no mutual interference. Within the working memory literature, such additive capacities could be argued to support separate storage of auditory-verbal and visuo-spatial material (e.g., Baddeley, 1986, 2012; Baddeley & Hitch, 1974; Baddeley & Logie, 1999; Logie, 2011), albeit that the task we used was IFR rather than ISR.

Second, we examined the preferred output orders in the recall of the dual-modality lists. The IFR task instructions gave participants the freedom to output the two types of stimuli in very different output orders if they so wished (e.g., recall the words in forward order starting with the first word, but initiate recall of the visuo-spatial lists with one of the last few dots), but this did not frequently occur. We had also anticipated that participants might output from one modality prior to recall from the other (cf., Broadbent, 1954; Dornbush, 1968; Margrain, 1967), but there was little or no evidence of within-modality clustering. Somewhat to our surprise, we found that the output orders in the recall of auditory-verbal and visuo-spatial stimuli were highly constrained. On the majority of trials, participants tended to initiate their recall of dual-modality lists with a word, and then they tended to transition alternately between dots and words from neighboring list positions (Howard & Kahana, 1999; Kahana, 1996). This highly constrained recall order makes it unlikely that participants were retrieving the dual-modality list items from separate modality-specific stores, but rather suggest that both dots and words were associated at encoding with a common temporal context which is used at test to guide retrieval. Many models of IFR assume that verbal items are associated with a temporal context (e.g., Davelaar et al., 2005; Glenberg & Swanson, 1986; Howard & Kahana, 2002a; Unsworth & Engle, 2007) or represented along a temporal dimension (Brown et al., 2007). Within the revised working memory model (Baddeley, 2000, 2012), our data could be seen as evidence supporting the need for separate modality-specific stores as well as the multi-modal episodic buffer, that is posited to integrate phonological, visual and other types of information.

Finally, we replicated and extended the findings observed by Cortis et al. (2015) who showed highly similar effects of list length on the serial position curves and output order in IFR of verbal and visuo-spatial stimuli. Like Cortis et al. (2015), we found bowed serial position curves for middle and longer lists of both modalities, and in these list lengths, the primacy and recency effects were greater in the IFR of spoken words than the IFR of visuo-spatial dots. Extending the findings of Cortis et al., we found there was only limited dual-task interference: recency was almost completely unaffected by the dual-modality presentation, but the magnitudes of the primacy effects were somewhat reduced, and this was more pronounced at list length 16. Like Cortis et al. (2015), we also found that participants tended to initiate IFR of short lists of auditory-verbal and visuo-spatial lists with the first list item, but as the list length increased, so participants tended to initiate IFR of longer lists of auditory-verbal and visuo-spatial lists with one of the last four list items. Extending the findings of Cortis et al., we found that the requirement to encode and recall two lists of stimuli simultaneously did not greatly affect the tendency to initiate recall with the first list item. This pattern is contrary to what one might expect if the tendency to initiate IFR of short lists of items with the first item was driven by a limited-capacity, domain-general short-term memory, but could again be explained if the auditory-verbal and visuo-spatial items were processed and maintained with very little mutual interference.

Experiment 2

Experiment 1 found that the IFR of dual-modality lists showed near-independent performance limitations in the recall of visuo-spatial dots and auditory-verbal words coupled with highly constrained output orders. Experiment 2 sought to replicate and to examine in more detail these somewhat surprising findings. First, we decided to obtain more comprehensive data sets of transitions with which we could perform fully conditionalized response probability analyses of the output orders. One difficulty in performing such analyses of the transitions with the data from Experiment 1 was that the data were distributed too thinly across five different list lengths. We circumvented this difficulty in Experiment 2 by collecting 50 trials of IFR data per participant using a single list length (list length 6).

Second, we were interested in knowing whether our findings of near-independent capacities and highly constrained output orders were somehow caused by the synchronized onsets of our dots and words in the dual-modality condition. The synchronized onsets may have encouraged participants to perceive the dual-modality lists as sequences of word-dot stimulus pairs, with the words and the dots highly associated, chunked, or bound together as integrated multi-modal stimuli or event files (e.g., Hommel, 2004). In Experiment 2, we sought to replicate Experiment 1, by comparing IFR of the single-modality lists with IFR of the dual-modality lists with synchronized onsets, and we sought to extend Experiment 1, by further comparing the IFR of dual-modality lists with alternating, non-contemporaneous, word-then-dot or dot-then-word onsets.

Thus, in Experiment 2, there were four groups of 20 participants who were each presented with 50 experimental lists of items for IFR. Participants in the words alone group heard 50 sequences of six words, presented via headphones, for IFR with spoken recall. Participants in the dots alone group were presented with 50 sequences of six visuo-spatial dots on the computer screen for IFR and recalled these locations using a computer mouse. Participants in the synchronized words and dots group were presented with 50 sequences of six spoken words and six visuo-spatial dots (with synchronized onsets) for IFR of both sets of stimuli, requiring spoken word responses and recall of the locations using a computer mouse. Finally, participants in the alternative words and dots group were presented with 50 sequences of six spoken words and six visuo-spatial dots, with the onsets of the words and dots alternating, so that the presentation of a word was not overlapping with the presentation of a dot. Half of the alternating sequences began with a dot; the remainder began with a word. We thought that this group of participants would be more likely to view the stimuli as an alternating, mixed modality list of twelve sequentially alternating stimuli than six integrated dual modality stimuli. Of interest was whether we would still see the essentially independent capacities and highly constrained output orders observed in the words and dots group when the list items were not contemporaneous.

Method

Participants. A total of 80 students from the University of Essex participated in exchange for either course credit or a small payment. None had participated in Experiment 1.

Materials. The materials were identical to those used in Experiment 1, except that we ensured that all spoken sound files lasted no longer than 600 ms, resulting in a decreased word pool of 332 words.

Design. The experiment used a mixed factorial design. There was one between-subjects independent variable, Stimulus group, with four levels: words only group, dots only group, synchronized words and dots group, and alternating words and dots group. There was one within-subjects independent variable, serial position with six levels: serial positions 1-6. The main dependent variable was the proportion of items correctly recalled.

Procedure. Each participant was randomly allocated into one of the four groups. All participants were informed that they would be shown two practice lists followed by two blocks each of 25 experimental trials. Participants were prompted with a “Ready?” prompt, and following a mouse click, the participants saw a sequence of stimuli for that trial. In the words only and the dots only groups, the stimuli were presented for 0.6s, with an inter-stimulus interval of 0.6s. In the synchronized words and dots group, participants saw 6 dots and heard 6 words and the onsets of the pairs of dots and words were synchronized. Each lasted 0.6s with 0.6s unfilled inter-stimulus interval. For the alternating words and dots group, the participants were presented with 6 dots and 6 words in an alternating schedule. Each item lasted 0.6s and on half of the trials, the sequences started with a word, whilst on the other half, the trials started with a dot. Recall was self-paced, and participants clicked a button to indicate that they had recalled all that they could remember. Participants were instructed to say out loud as many words as they could remember in any order, and click on the screen in any order to indicate the spatial locations of as many of the dots, as they could remember. In all groups, participants pressed the “submit” button once they were confident that they had recalled all that they could remember, and this initiated the next trial.

Results

Overall Accuracy. Figure 5 shows the accuracy in recalling the spoken words and the visuo-spatial dots stimuli for each of the four groups. As in Experiment 1, a word was scored as correctly recalled only if it was identical to the word that was presented. A strict scoring criterion for visuo-spatial responses was also used: a clicked location was scored as correctly recalled only if it fell within the circumference of a presented dot. We first analyzed performance for the different groups within stimulus domains, before comparing the accuracy across modalities.

--Figure 5 about here--

Considering first the recall of the spoken words across the different groups in Figure 5A, a between subjects ANOVA (group: words alone, synchronized words and dots, alternating words and dots) revealed that there was a non-significant main effect of group on verbal accuracy, $F(2,57) = .403$, $MSE = .007$, $\eta^2_p = .014$, $p = .670$.

Considering next the recall of the visuo-spatial dots across the different groups in Figure 5B, a between subjects ANOVA (group: dots alone, words and dots with synchronized onsets, alternating words and dots) revealed that there was a significant main effect of group on non-verbal accuracy, $F(2,57) = 7.98$, $MSE = .005$, $\eta^2_p = .219$, $p = .001$. Post-hoc Bonferroni tests showed that there was a significant reduction in non-verbal accuracy in the alternating group relative to the dots alone group ($p = .001$); all other tests were non-significant (all *p*s > .064). It should again be noted that participants in the dots alone group made significantly more mouse clicks (5,884) than in the synchronized words and dots onset group (5,043), and the alternating words and dots group (4,753), $F(2,57) = 15.5$, $MSE = .012$, $\eta^2_p = .353$, $p < .001$, suggesting that at least some of the numeric advantage for the dots only condition may reflect increased tendency to make additional mouse clicks.

Finally, two analyses were performed comparing the overall accuracy in recalling the words and the dots. An independent-samples *t*-test compared performance in the words alone group with performance in the dots alone group, and this revealed an auditory-verbal advantage in the single-modality groups, $t(38) = 12.61$; $p < .001$; $d = 3.99$. A 2 (group: synchronized words and dots, alternating words and dots) x 2 (stimulus type: words, dots) mixed ANOVA, compared the overall accuracy in recalling the words and the dots in the dual –modality conditions. This showed that there was a significant main effect of stimulus type, $F(1,38) = 346.4$, $MSE = .006$, $\eta^2_p = .901$, $p < .001$, with words recalled better than dots, a non-significant main effect of group, $F(1,38) = .320$, $MSE = .007$, $\eta^2_p = .008$, $p = .575$, and a non-significant interaction on overall accuracy, $F(1,38) = 2.019$, $MSE = .006$, $\eta^2_p = .050$, $p = .164$.

Overall, there is clear evidence that (1) participants recalled the words more accurately than they recalled the dots in both the single-modality and the dual-modality conditions, that (2) recalling the words was not greatly affected by the requirement to also recall dots in either of the two dual-modality groups, whereas (3) recalling the dots was somewhat affected by the requirement to also recall words, and the decrease in the recall of the dots reached significance in the comparison with the alternating words and dots condition.

Serial position curves. Figure 6 show the serial position curves for each of the four groups. Figure 6A shows the serial position curves for the auditory-verbal data. A 3 (group: words alone, synchronized words and dots, alternating words and dots) x 6 (serial position) mixed ANOVA revealed that there was a non-significant main effect of group, $F(2, 57) = 0.403$, $MSE = .007$, $\eta^2_p = .014$, $p = .670$, a significant main effect of serial position, $F(5, 285) = 14.9$, $MSE = .027$, $\eta^2_p = .207$, $p < .001$, and a non-significant interaction, $F(10, 285) = 1.45$, $MSE = .027$, $\eta^2_p = .048$, $p = .158$.

There was evidence of bowed serial position curve in all three verbal serial position curves, but little effect of a dual-task reduction in recalling the words.

--Figure 6 about here --

Figure 6B shows the serial position curves for the visuo-spatial dot data. A 3 (group: dots alone, synchronized words and dots, alternating words and dots) x 6 (serial position) mixed ANOVA revealed that there was a significant main effect of group, $F(2, 57) = 7.98$, $MSE = .030$, $\eta^2_p = .219$, $p = .001$, a significant main effect of serial position, $F(5, 285) = 11.5$, $MSE = .015$, $\eta^2_p = .219$, $p < .001$, and a significant interaction, $F(10, 285) = 2.37$, $MSE = .015$, $\eta^2_p = .077$, $p = .010$. Pairwise comparison (with Bonferroni corrections) of the main effect of group showed that the only significant difference between the groups was that the recall in the dots only group was significantly greater than in the alternating presentation group ($p = .001$). The significant interaction between group and serial position was the result of a reduced recency effect (serial position 4 onwards) in the alternating words and dot group. However, overall there was evidence of bowed serial position curve in all three verbal serial position curves, and evidence of a dual task reduction in recalling the dots in the alternating group.

Probability of First Recall Data. Table 3 shows the number of trials in which participants started their recall with each of the six serial positions. Across both word and dot stimulus modalities, participants were more likely to initiate recall with the first item in the single-modality conditions and the alternating words and dots condition. By contrast, in the synchronized words and dots condition, participants were more likely to initiate recall from both stimulus modalities with the last item on the list.

--Table 3 about here--

Detailed Analysis of Output Order in the dual-modality conditions. In this section, we again consider in great detail the output orders in the two dual-modality groups, considering the order of outputs both within- and across- modalities. Consistent with Experiment 1, there are four main findings.

First, in both the dual-modality groups, participants preferred to initiate recall with a word. This word-first tendency occurred on 87% of trials in the synchronized words and dots group, and this word-first tendency was also present in the alternating words and dots group (78%), albeit at a significantly reduced rate, $F(1,38) = 4.21$, $MSE = .019$, $\eta^2_p = .100$, $p = .047$. Interestingly, this word-first tendency was present when the dots in the alternating words and dots condition were presented first (76.2%) as well as when the words were presented first (79.8%), and this difference was not significant, $F(1,19) = 1.45$, $MSE = .009$, $\eta^2_p = .071$, $p = .243$.

Second, in both the dual-modality groups, there was a tight coupling on any given trial between the first word recalled and the first dot recalled. Table 4 shows the distribution of first responses at each serial position for the auditory-verbal and visuo-spatial modalities for the synchronized onsets group and the alternating words and dots group, respectively. In the synchronized onsets group there were 601 trials where participants made correct first responses in both modalities, out of which 515 (85.7%) had matching within-modality serial positions (e.g., starting with the first word and first dot respectively). In the alternating words and dots group, there were 467 such trials, of which 357 (76.4%) had matching within-modality output orders.

There were however, some subtle differences between the two dual-modality groups regarding which serial position participants preferred to start with. Consistent with the differences noted in Table 3, participants in the synchronized group were most likely to start with the last item of each modality on 28.9% of trials, whereas participants in the alternating presentation condition were more like to start with serial position 1 in both modalities (21.9%) and initiate recall with the last list item of both modalities on only 8.3% of trials.

--Table 4 about here--

Third, in both dual-modality groups, participants tended to alternate between outputting words and dots. For the synchronized words and dots group, of the 3,127 recorded transitions from list length 6 (i.e., six words and six dots), the majority (2,440, 78.0%) were between items of different modalities. Specifically, there were 1,352 (43.2%) transitions from words to dots (WD transition) and 1,088 (34.8%) transitions from dot to word (DW transitions). In contrast, only a minority of transitions were between items from the same modality: 549 (17.6%) transitions were between two words (WW transitions) and 138 (4.4%) transitions were between two dots (DD transitions). Similarly, for the alternating words and dots group: of the 2,762 recorded transitions from list length 6, the majority (2,169, 78.5%) were between items of different modalities. Specifically, there were 1,147 (41.5%) transitions from words to dots (WD transitions) and 1,022 (37.0%) transitions from dots to words (DW transitions). Again, only a minority of transitions were between items from the same modality: 487 (17.6%) were between two words (WW transitions), and only 106 (3.8%) were between two dots (DD transitions).

Finally, the output orders in both dual-modality groups were highly constrained. Due to the outputs from 50 trials of each list length, we were able to plot how the conditionalized response probabilities (CRPs) varied with different lag transitions. For each participant, we summed the total number of actual transitions that were made between stimuli of different modalities and between items presented at different serial positions (the transition distance measured as the lag, serial position of output $n+1$ - serial position of output n). For each participant, we also summed the total number of opportunities that there were to make these transitions, by considering at each point in the output the modality and the serial position of all the stimuli that had yet to be recalled. The CRP for each type of transition for each participant was calculated by dividing the number of observed transitions by the total number of possible transitions

As Figure 7A shows, when participants were asked to recall the synchronized words and dots in any order, they were most likely to recall a word followed by the dot of the same within-modality serial position (WD lag 0), and transition from dots to neighboring words (DW lag -1, 0, +1). Figure 7B shows that transitions in the alternating words and dot condition were also highly constrained. In the alternating words and dot conditions, there were effectively 12 different serial positions, such that lags can range from -11 to +11 with no lag 0. In the Alternating words and dots group, participants were most likely to transition from a word to one of the nearest dots of the same within-modality serial position (WD lag +1 or -1), and transition from a dot to a nearby word (DW lag +1, +3, or -1).

--Figure 7 about here--

Discussion

Experiment 2 aimed to further examine the near-independent recall performance limits and the highly constrained output orders that were observed in the synchronized dual-modality conditions in Experiment 1. This was achieved by collecting an extensive set of data of 50 trials per participant using a single within-modality list length of six items. As in Experiment 1, we contrasted the performance of the synchronized words and dots group (who encoded and recalled stimuli from both modalities) with performance of the words alone group and the dots alone group (who encoded and recalled stimuli from only one of the two modalities). We introduced a fourth group, the alternating words and dots group, to determine whether or not the synchronized onsets were responsible for the near-independent recall performance limits and the highly constrained output orders that were observed in Experiment 1.

Regarding the capacity or performance limits in recalling auditory-verbal and visuo-spatial stimuli, Figure 5 illustrates that there was very little difference between the proportion of words recalled in the three verbal recall groups: words alone group, the synchronized words and dots group and the alternating words and dots group. Participants who were asked to encode and recall a set of 6 words and a set of 6 dots recalled just as many words as those participants who were asked to encode and recall only a set of 6 words.

However, there was some evidence of a decline in the recall of the dots across the three different dot groups. Recall in the dots alone condition was significantly different from the recall of the dots in the alternating words and dots group. Consistent with Experiment 1, the recall in the dots alone condition did not differ significantly from the recall of the dots in the synchronized words and dots group. However, it should be noted that there was also no significant difference between the recall of the dots in the two dual-modality groups, so it seems prudent to assume that there might be some, relatively small degree of interference between the words and the dots when trying to encode and recall from both modalities. Although the dual-task decrement in the visuo-spatial recall was significant for the alternating words and dots group, it is worth emphasizing that even in this comparison, the absolute magnitude of the decrement was relatively small. Participants in the dual-modality conditions were able to recall over 75% of the dots recalled in the dots alone group whilst still recalling over 95% of the words recalled in the words alone group. We argue that this represents near-independent performance limits in performing IFR of dual-modality lists. If anything, there is evidence of an asymmetry in the patterns of cross-modal interference: a concurrent visuo-spatial task appears to disrupt an auditory-verbal task to a lesser extent than a concurrent auditory-verbal task appears disrupts a visuo-spatial task, a finding consistent with the visual-spatial asymmetry in serial recall (e.g., Cowan & Morey, 2007; Morey & Mall, 2012; Morey & Miron, 2016; Morey, et al., 2013).

Consistent with Experiment 1 and Cortis et al. (2015), participants correctly recalled a higher proportion of auditory-verbal words than the visuo-spatial dots, and the serial position curves showed more pronounced primacy and recency with words than dots. There was also a greater tendency in the auditory-verbal PFR data for a more graded and more extended recency over the last few items than for the visuo-spatial PFR data.

Regarding the output orders, participants in both dual-modality groups tended to show highly constrained output orders. Consistent with Experiment 1, they tended to initiate recall with a word rather than a dot, they tended to alternate between modalities in successive responses, and they tended to recall successive items from either the same or neighboring serial positions. Although these tendencies were slightly reduced in the alternating words and dot group, there were clear and consistent preferences for making successive responses from neighboring items from different modalities. We argue that these highly constrained, alternating output orders suggest that the retrieval of the visuo-spatial dots and the auditory-verbal words is unlikely to be from separate, independent modality-specific stores with independent retrieval mechanisms.

The fact that the highly constrained cross-modal output order was present even in the words and dots with alternating stimulus onsets group additionally shows that the finding does not necessitate a specialized explanation based on the chunking, association or binding of stimuli with synchronized onsets. Rather, we argue that the visual dots and spoken words were associated with a common time- or position-sensitive episodic memory that was used at test to guide retrieval of both words and dots. The lack of catastrophic dual-task decrement could then be explained if one assumed that the representations of the associated dots and words were sufficiently different from each other for there to be very little mutual interference. Alternatively, a more extreme explanation for the lack of catastrophic dual-task decrement would be to assume that the dots and the words are encoded in separate modality-specific stores but are integrated within a common temporal- or position-sensitive episodic buffer (Baddeley, 2000, 2012).

Experiment 3

Experiments 1 and 2 had found evidence for near-independent IFR performance limitations and highly constrained output orders for synchronized word and dot lists of stimuli relative to lists of single-modality items. Experiment 2 decoupled the onsets of the word and the dot stimuli by alternating the timing of the word and dots. However, this manipulation did not greatly reduce recall, or the tendency for participants to alternate between successive word and dot responses from neighboring list items.

Although lists in the alternating words and dots group had decoupled the onsets of the two types of stimuli, it was still possible that participants' near-independent performance or capacity limits and highly constrained outputs were the result of a highly regular, highly predictable sequence structure. One possibility was that the higher order list structure in the synchronized and alternating lists encouraged the grouping of the dual-modality lists into sequences of cross-modal pairs of items rather than two independent streams of single-modality lists. It is recognized in serial recall that the relative timings and orders of constituent list elements can influence the grouping or streaming structure perceived by participants (e.g., Hughes, Marsh & Jones, 2009; Jones, 1993; Nicholls & Jones, 2002), which can in turn affect subsequent recall. A second possibility was that with predictable list structure, participants could actively integrate the words with the dots, perhaps by thinking of the object associated with each word and then imagining that item in the paired dot location (a form of cross-modal binding, e.g., Allen, Hitch & Baddeley, 2009). This active encoding strategy could occur for pairs of words and dots presented with synchronized onsets, but could also be used with predictable alternating word dot or dot word pairs, if one assumed that one or other of the stimuli could be easily maintained for 0.6 s until the arrival of the paired item. Thus, it may be that the dual-modality findings in IFR observed in Experiments 1 and 2 represent special cases using highly structured lists, and that our dual-modality IFR findings may not generalize to less predictable list structures.

Thus, the main aim of Experiment 3 was to determine whether the near-independent performance limitations and highly constrained output orders in the IFR of dual-modality lists that were observed in Experiments 1 and 2 would generalize to lists of 6 words and 6 dots that were presented sequentially in a random order. If the tendency to alternate response modalities was simply due to the list structure at presentation, then there should be a greatly reduced number of alternations in this experiment. There is some evidence (e.g. Murdock & Walker, 1969) that when participants are presented with mixed-modality lists of auditory and visual words presented in a random order, participants tend to cluster their outputs by input modality (i.e. channel by channel output; also see Murdock & Carey, 1972; Nilsson, Ohlsson, & Rönnberg, 1977). They also found the typical modality effect, whereby auditory-words were better recalled than visual words and that this advantage was seen in the recency portion of the serial position curve. Moreover, when presenting participants with mixed-modality lists for ISR, Greene (1989) found that the recall advantage for auditory items in mixed-modality lists was not found solely for the last few items but across all serial positions. He suggested that auditory items were recalled better since these have two dimensions that make them highly accessible: one that relates to the item's identity and the other that relates to acoustic properties, such as whether the voice was that of a male or female, or whether the voice was loud or quiet.

Thus, Experiment 3 compared the performance limitations and the output orders of three groups of participants who performed IFR on single-modality lists of six dots alone, lists of six spoken words, or on the random dual-modality lists composed of randomly-ordered sequences of six visuo-spatial dots and six auditory-verbal words.

Method

Participants. A total number of 60 students from the University of Essex participated in exchange for either course credit or a small payment.

Materials. The materials were identical to those used in Experiment 2.

Design. The experiment used a mixed factorial design. There was one between-subjects independent variable: stimulus group with three levels: the words only group, the dots only group, and the randomized words and dots group. Overall, there was one within-subjects independent variable: serial position with six levels for the auditory-verbal and visuo-spatial group and twelve levels in the randomized words and dots group. The main dependent variable was the proportion of items correctly recalled.

Procedure. Each participant was randomly allocated upon arrival into one of the three groups, such that each group consisted of 20 participants. Participants in all three groups received two practice lists of six stimuli (six words and six dots in the randomized presentation group), followed by 50 experimental lists of stimuli separated into two blocks. Each trial was preceded by a "Ready?" prompt, instructing participants to initiate the next trial with a button press or mouse click when they were ready to do so

For all three groups, the trial structure of every trial was created by randomly assigning six auditory words and six visuo-spatial dot locations into a 12-item sequence with onsets that were separated by 0.60s, with no inter-stimulus interval. Using these list structures, participants in the words only group were presented only with the spoken words via headphones, and the screen remained blank during the intervals that had been set aside for the dots. Similarly, participants in the dots only group were presented only with the visuo-spatial dot locations, and there was silence and a blank screen during the intervals that had been set aside for the words. Finally, participants in the randomized words and dots group were presented with both the visuo-spatial dots and the spoken words (via headphones) in the randomly assigned list positions. In all groups, participants pressed the "submit" button once they were confident that they have submitted all the responses they could recall and this initiated the next trial.

Results

Overall Accuracy. Figure 7B shows the accuracy in recalling the spoken words and the visuo-spatial dots stimuli for each of the three groups. As in Experiment 1 and 2, a word was scored as correctly recalled only if it was identical to the word that was presented. A strict scoring criterion for visuo-spatial responses was also used: a clicked location was scored as correctly recalled only if it fell within the circumference of a presented dot.

Considering first the recall of the spoken words across the different groups, an independent samples t-test comparing the proportion of words recalled in the words alone group with the proportion of words recalled in the randomized words and dots group showed that there was a non-significant difference in verbal accuracy between the two groups, $t(1,38) = 0.30$, $p = .769$. Considering next the recall of the visuo-spatial dots across the different groups, an independent samples t-test comparing the proportion of dots recalled in the dots alone group with the proportion of dots recalled in the randomized words and dots group showed that there was a significant difference in visuo-spatial performance, $t(38) = 3.51$, $p < .001$. Thus, when participants were required to remember stimuli in two stimulus modalities rather than just one modality, the recall of the visuo-spatial dot locations dropped from .32 to .25 whereas the recall of the spoken words dropped from .64 to .63.

Finally, two analyses were performed comparing the recall of the words with the recall of the dots. Considering first the between-group comparison between the single-modality groups, an independent samples t-test confirmed that a greater proportion of words in the words alone group were recalled than dots in the dots alone group, $t\ (38) = 12.83$, $p < .001$. Considering next the within-group comparison between the words and the dots recalled in the randomized words and dots group, a paired samples t-test confirmed that a greater proportion of words were recalled than dots in the randomized words and dots group, $t(19) = 18.33$, $p < .001$.

Serial Position Curves. Figure 8 shows the serial position curves for each of the three groups for the verbal and visuo-spatial modalities. Considering first the recall of the words, a 2 (group: words alone group and the randomized words and dots group) x 6 (serial position) mixed ANOVA on the verbal data showed that there was a non-significant main effect of group, $F(1,38) = .087$, $MSE = .050$, $\eta^2_p = .002$, $p = .769$, a significant main effect of serial position, $F(5,190) = 39.3$, $MSE = .013$, $\eta^2_p = .508$, $p < .001$, and a non-significant interaction, $F(5,190) = 1.19$, $MSE = .013$, $\eta^2_p = .030$, $p = .315$. Overall, the serial position curves for the verbal stimuli are very similar in shape, with both primacy and recency effects, leading to a bowed serial position curve.

--Figure 8 about here--

Considering next the recall of the visuo-spatial dot locations, a 2 (group: dots alone group and the randomized words and dots group) x 6 (serial position) mixed ANOVA showed that there was a significant main effect of group, $F(1,38) = 12.3$, $MSE = .023$, $\eta^2_p = .244$, $p = .001$, a significant main effect of serial position, $F(5,190) = 10.4$, $MSE = .008$, $\eta^2_p = .215$, $p < .001$, and a non-significant interaction, $F(5,190) = .551$, $MSE = .008$, $\eta^2_p = .014$, $p = .737$. The serial position curves for the visuo-spatial stimuli also show some bowing but it is clear that requirement to also encode and recall the words in the randomized presentation group has resulted in equivalent reduction across all serial positions.

Probability of First Recall Data. Table 5 shows the number of trials in which participants started their recall with each of the six serial positions for each modality across the three groups. Similar to the alternating words and dots group in Experiment 2, participants in Experiment 3 were more likely to initiate their recall with the first list item in both the single- and the dual-modality groups across both the verbal and visuo-spatial modalities.

--Table 5 about here--

Detailed Analysis of Output Order in the dual-modality conditions. Consistent with the previous experiments, participants initiated output in the randomized words and dots group with a word (71.2% of trials) more often than a dot (28.8%). A within-subjects ANOVA confirmed that the proportion of trials in which recall was initiated with a word was not significantly affected by whether the first presented item was a word or a dot, $F(1,19) = 2.57$, $MSE = .009$, $\eta^2_p = .119$, $p = .126$.

Table 6 shows the distribution of recalled serial positions that were first recalled for the verbal and visuo-spatial modality of the randomized presentation group, to determine whether first responses across the two modalities were coupled or were independent. There were 461 trials where participants made correct first responses in both modalities, out of which 220 (47.7%) responses had matching within-modality serial positions. These data are more similar to the alternating presentation group than the synchronized onsets group in Experiment 2, in that participants were more likely to start recall at the start of the list than with the last item (17.4% versus 3.2% of all responses). However, the tendency to start each modality with matching serial positions seems somewhat reduced in the present experiment relative to the data in Experiment 2, and this is seen with smaller values across the leading diagonal.

--Table 6 about here--

Although participants were fairly likely to alternate between outputting words and dots when presented with a randomized sequence of auditory-verbal words and visuo-spatial dots, this was reduced relative to the previous experiments, whereas the tendency to output consecutively two words was relatively increased. Out of the 2,704 transitions, there was a considerable number that were between different modalities: 793 were from dot-to-word (DW) transitions and 861 were from word to dot (WD) transitions. As in Experiments 1 and 2, the number of dot-to-dot (DD) transitions was low (142), but contrary to the prior experiments, there was a high tendency for word-to-word transitions (908). Of course, with randomized list structures, it is not immediately clear how these transitions reflect the structure of the lists. Figure 9 shows the Lag-CRP functions plotting the different transitions for the randomized word and dots group. These functions reflect both the frequency of observed transitions and also the different opportunities to make these different transitions over different lags.

--Figure 9 about here--

It can be seen that participants were more likely to recall consecutive items from temporally close serial positions, resulting in heightened frequency of smaller lags. Furthermore, +1 and +2 lags are most likely to be WW transitions and whereas there is a marked asymmetric lag recency effect in same-modality transitions, this is not as pronounced in the cross-modality transitions. Whereas the tendency to go in backwards order is markedly reduced for within-modality transitions, participants tend often to make backward transitions order across modalities.

Discussion

Experiment 3 examined the accuracy and output order in randomized lists of words and dots in order to evaluate whether the patterns of near-independent performance limitations and highly constrained output orders observed in Experiments 1 and 2 were because these earlier experiments used highly predictable (synchronized or alternating) input structures. There were two main sets of findings.

We consider first the accuracy in the IFR across the three groups. Consistent with Experiments 1 and 2, it is again clear that the words were recalled more accurately than the dots. The accuracy of the Randomized words and dots group was similar to the Alternating words and dots group of Experiment 2. Consistent with the recall of the alternating lists, visuo-spatial IFR did not significantly disrupt auditory-verbal IFR, but a concurrent auditory-verbal IFR did significantly reduce visuo-spatial IFR performance. This asymmetric pattern of cross-modal interference is consistent with the visual-spatial asymmetry often observed in serial recall (e.g., Cowan & Morey, 2007; Morey & Mall, 2012; Morey & Miron, 2016; Morey, et al., 2013). It should be noted that although the recall of the dots was reduced significantly in the dual-task conditions, participants in the randomized words and dots group could still recall over 98% of the words recalled by the words alone group whilst maintaining over 68% of the visuo-spatial dots alone group. It would appear that a highly predictable output order is not necessary for the highly accurate recall of the words and the good recall of the dots in dual-modality IFR.

We consider next the patterns of output orders in the IFR across the three groups. It is clear that in both stimulus modalities, participants tended to initiate their recall with the first item, or to a lesser extent with one of the last items on the list (Cortis et al., 2015; Ward et al., 2010). Consistent with Experiments 1 and 2, participants tended to initiate recall with a word more frequently than they to initiated recall with a dot. The tendency to have highly constrained output orders was present but rather reduced with randomized presentations. The initial recall in the two modalities showed mutual tendencies to start with the first list item in the two modalities, but there was little evidence of coupled initial recall from middle list items. The subsequent outputs were highly constrained: participants tended to recall items that were temporally contiguous regardless of modality and participants were highly willing to make cross-modality transitions. However, there was also greater evidence of participants in the randomized groups making more word-to-word transitions than in our previous dual-modality groups.

Finally, the findings of Experiment 3 confirmed several similarities and differences between the IFR of words and dots. As noted in Experiments 1 and 2, there were bowed serial position curves with both types of stimuli, but the degree of primacy and recency was more pronounced with words than dots. Participants were generally more accurate in recalling the words than the dots, although this was to some extent a result of the strict criterion for recalling a dot (see next section: On the issue of low visuo-spatial accuracy). Participants also tended to initiate recall with either the first or one of the last few list items with both types of stimuli, but the tendency to initiate recall with a recency item was more graded and more extended with words, whereas the tendency to initiate recall with a recency item was more limited to the final list items with dots. We will return to the similarities and differences between the IFR of auditory-verbal and visuo-spatial stimuli in the General Discussion.

On the issue of low visuo-spatial accuracy

In all our analyses on visuo-spatial accuracy, a strict criterion has been used in which a mouse click was scored as correct only if the location of the mouse click fell within the circumference of a target dot. It was necessary to adopt this rather strict criterion in Experiment 1 because in that experiment there were up to 16 possible dot stimuli on a trial. Since each 35mm diameter dot occupied approximately 0.9% of the total 375mm x 280mm area, even the strict criterion led to a guessing rate of around 15% for a random first response to a 16-dot list. Adopting a more lenient scoring, by increasing the allowable target zone to twice the stimulus radius (a circle of 70mm diameter), would help give credit to near misses, but each such target area would then occupy 3.7% of the total area of the screen, such that in the case of a 16-dot list, an initial random response would be more likely than not to be scored as correct (58.6%).

However, because Experiment 2 used only 6-item lists, it would be more reasonable to increase the size of the target zone, perhaps to as much as a 70mm circle. Using a strict criterion, guessing with only 6 dots is less than 5.5%; relaxing the criterion to a 70 mm circle would result in an initial random mouse click being scored as correct on around 22% of occasions, which is not entirely unacceptable. Consequently, in this penultimate section of this paper we discuss the implications of using a more lenient visuo-spatial scoring on the performance limits and the output order for the 6-item lists used in Experiments 2 and 3.

Two lines of evidence suggest that we should embrace a more lenient scoring system. First, the mean proportion of correctly recalled visuo-spatial dots in the dots alone group of Experiments 2 and 3 were .37 and .32, respectively, considerably lower than the mean proportion of correctly recalled words in the words alone groups (.64 and .64, respectively). Second, Tables 3 and 5 showed that 41.5% and 51.0% of the first visuo-spatial response of Experiment 2 and 3, respectively, were scored as either incorrect or non-responses. Consequently, we were concerned that we had excluded from our analyses a potentially large number of near-misses, and that as a result, interesting patterns in the data may have been obscured.

The analyses in this section re-plotted the key aspects of the data in Experiments 2 and 3, but using both the strict and four increasingly lenient criteria. Therefore, the target radius of visuo-spatial target location was widened from the strict criterion (x1 stimulus radius) to radii of magnitude x1.25, x1.50, x1.75, and x2 resulting in acceptable target diameters of 35mm, 43.8mm, 52.5mm, 61.2mm, and 70mm, respectively. Figure 10A and Figure 10B show that increasing the leniency criteria greatly reduced the number of visuo-spatial errors in the data from Experiments 2 and 3, respectively. By adjusting the criterion for accuracy, to twice the dot stimulus (x2), we have successfully accounted for nearly half of all dot errors, such that we should be able to see patterns of IFR performance that might otherwise have been obscured using our strict criterion.

Figure 10C and Figure 10D show that dot accuracy approaches auditory-verbal accuracy using more lenient criteria (between .62 and .64). Additionally, when using the more lenient criteria, the difference between the single-and dual-modality IFR performance somewhat increased. Although the differences between the dots only and the dots in the dual-modality IFR groups did not always reach significance with the x1 strict criterion, when the data are reanalysed using the more lenient x2 criterion, a significant difference emerges between all dots only and dual-modality lists. Thus, in Experiment 2, even the difference between the dots only group and the synchronized words and dots group, became significant with the x2 criterion, $t(38) = 3.81$, $p < .001$. It should be noted that even at the x2 criterion there was not a significant difference in the recall of the dots from the synchronized words and dots and the alternating words and dots group, $t(38) = 1.29$, $p = .20$. Finally, it is worth reflecting that the magnitude of the dual-task decrements remains approximately constant as the criterion is relaxed. In Experiment 2, the recall of the dots from the synchronized words and dots group as a proportion of the dots recalled in the dots alone group varied from .86, .84, .84, .84, .84 (from x1 through to x2). Similarly, the recall of the dots from the alternating words and dots group in Experiment 2 as a proportion of the dots recalled in the dots alone group varied from .76, .77, .76, .77, .78 (from x1 through to x2). Finally, the recall of the dots from the randomized words and dots group in Experiment 3 as a proportion of the dots recalled in the dots alone group varied from .79, .77, .76, .77, .79 (from x1 through to x2).

--Figure 10 about here--

Finally, Figure 10E and Figure 10F shows the proportion of correct initial responses in which the first responses in Experiments 2 and 3 were highly constrained; that is where the serial position of the first verbal response was the same as the serial position of the first visuo-spatial response. There are two main points from these two panels. First, it is clear that the vast majority of the first outputs are constrained in the synchronized and alternating groups of Experiment 2, but the degree of tight coupling between the first responses from each modality was greatly reduced in the randomized condition of Experiment 3. Second, as the near horizontal lines in Figure 10E and Figure 10F indicate, these findings are largely unaffected by increasing the leniency of the visuo-spatial responses. Overall, Figure 10 shows that even when using a more lenient criteria for visuo-spatial dot accuracy, we find that participants in the different dual-modality groups of Experiments 2 and 3 can recall almost all of the words and between .78 and .84 of the dots that are recalled in the single-modality groups, and show just as constrained initial outputs as when using the strict criteria.

In order to determine whether the change in dot scoring criteria affected the shapes of the serial position curves, we re-plotted the curves for the visuo-spatial IFR data of Experiments 2 and 3 using the most lenient x2 scoring criterion. Figure 11A shows that there continue to be bowed serial position curves for the dot stimuli in Experiment 2 and Figure 11B shows that there continue to be bowed serial position curves for the dot stimuli in Experiment 3. In both experiments, the degree of primacy and recency remains reduced relative to the IFR of the words.

--Figure 11 about here--

A final set of analyses were performed on the Lag-CRP functions in the dual-modality conditions of Experiments 2 and 3 using the lenient x2 visuo-spatial scoring criteria. Figure 12A show the Lag-CRP functions for the Synchronized words and dots group of Experiment 2, Figure 12B show the Lag-CRP functions for the Alternating words and dots group of Experiment 2, and Figure 12C show the Lag-CRP functions for the Randomized words and dots group of Experiment 3. As can be seen, relaxing the scoring criterion did not greatly affect the shapes of the Lag-CRP functions. There remained very high probabilities of cross-modality near-neighbor transitions in the two dual-modality conditions of Experiment 3, whereas there continued to be both cross-modality near-neighbor transitions as well as higher word-to-word transitions in the Randomized words and dots group of Experiment 3.

--Figure 11 about here--

General Discussion

Over three experiments, we have sought to address three main research questions: (1) Do verbal immediate memory and visuo-spatial immediate memory share a common performance limit or do they exhibit independent or near-independent performance or capacity limits? (2) Are the output orders from verbal and visuo-spatial immediate memory essentially independent or highly constrained? and (3) What are the similarities and differences between the serial position curves and output orders in the IFR of lists of auditory-verbal and visuo-spatial items?

IFR Accuracy: Performance limits in single- and dual-modality lists

We examined the IFR performance limits in auditory-verbal and visuo-spatial immediate memory, by comparing the proportion of items recalled in the IFR of single-modality lists with the proportion of items recalled in the IFR of dual-modality lists. In three experiments, we presented spoken lists of words via headphones for participants to recall orally and / or we presented visuo-spatial sequences of dot circles on a computer screen for participants to recall by clicking with a computer mouse. Experiment 1 varied the list length from 1, 2, 4, 8 or 16 within-modality list items and used synchronized onsets of words and dots in the dual-modality groups. Experiments 2 and 3 used trials with only 6 within-modality list items, and contrasted single-modality lists with synchronized words and dots (Experiment 2), alternating words and dots (Experiment 2), and randomized words and dots (Experiment 3).

Considering first the performance limitations for auditory-verbal IFR, it is clear from the data from all three of our experiments that the proportions of words recalled from the dual-modality lists of words and dots were not significantly different from the proportions of words recalled from the corresponding single-modality, words alone group. Thus, in Experiment 1, the recall of words from the single-modality and synchronized dual-modality lists were .72 and .72; in Experiment 2, the recall of words from the single-modality and the synchronized and alternating dual-modality lists were .64, .62, and .63, respectively; whereas in Experiment 3, the recall of words from the single-modality and randomized dual-modality lists were .64 and .63. We can therefore conclude with reasonable confidence that the auditory-verbal IFR performance limits are relatively unaffected by an additional, simultaneous requirement to perform visual-spatial IFR. Our analyses show that participants in dual-modality IFR lists can recall over 95% of the words recalled in the single-modality, words only conditions.

Considering next the performance limitations in visuo-spatial IFR, our data suggest that the proportions of dots recalled from the dual-modality lists of words and dots were somewhat lower than the proportions of dots recalled from the corresponding single-modality, dots alone lists. Thus, in Experiment 1, the recall of dots from the single-modality and synchronized dual-modality lists were .51 and .48; in Experiment 2, the recall of dots from the single-modality, and the synchronized and alternating dual-modality lists were .37, .31, and .28, respectively; whereas in Experiment 3, the recall of dots from the single-modality and the randomized dual-modality lists were .37 and .25. The dual-task decrements reached significance with strict scoring with the alternating words and dots group of Experiment 2 and the randomized words and dots group of Experiment 3, but using the strict scoring criterion the dual-task decrement was not significant for the synchronized words and dots groups in Experiments 1 and 2. Two sets of comparison suggest that there are nevertheless significant dual-task decrements more generally with the visuo-spatial IFR. First, we found no significant difference in the recall of the dots between the different dual-modality conditions and second, significant dual-task decrement was found even with synchronized words and dots in Experiment 2 when a more lenient scoring system was used. Our analyses show that participants in dual-modality IFR lists can recall between 74% an 85% of the dots recalled in the single-modality, dots only conditions.

Therefore, our dual-modality IFR findings differ considerably from classic verbal list learning studies examining the IFR of words presented at different rates that *have kept the total duration of the list constant* (e.g., Roberts, 1972; Waugh, 1967). These studies have shown only a very small increase (Roberts, 1972) or no increase (Waugh, 1967) in the number of items recalled with increasing to-be-remembered items. Even when the total presentation time of a list is increased in line with the number of items to-be-remembered, classic studies on the list length effects have shown that as the number of items in the list increases, so the number of items recalled increases at a greatly reduced proportion correct (Murdock, 1962; Ward, 2002, Ward et al., 2010). Thus, our findings suggest that the auditory-verbal words and the visuo-spatial dots are not encoded as equivalent multimodal or amodal list items, but are encoded as items from two different stimulus domains in a near-independent manner.

Of course, our methodology differs from prior studies of list length in that it used both verbal and non-verbal stimuli, and it is perhaps not surprising that words and dots produce very little mutual interference. However, it is telling that theories of IFR have been dominated by explaining the recall of word lists, with few, if any, contemporary theories of IFR explicitly considering the recall of non-verbal stimuli. By contrast, theories of ISR (and theories of working memory, more generally) have paid great attention to the stimulus modality in which the items were presented. Most relevant is the cross-modal asymmetry in the pattern of dual-task performance in ISR. ISR of auditory-verbal lists is relatively unaffected by concurrent visuo-spatial processing, but ISR of visuo-spatial lists is significantly affected by concurrent auditory-verbal processing (e.g., Cowan & Morey, 2007; Morey & Mall, 2012; Morey & Miron, 2016; Morey, et al., 2013). This asymmetric pattern of dual-task interference observed in ISR closely resembles our own data from IFR: when participants were required to simultaneously perform IFR of auditory-verbal and visuo-spatial lists, IFR of auditory-verbal lists was essentially unaffected by concurrent visuo-spatial IFR, but IFR of visuo-spatial lists was significantly affected by concurrent auditory-verbal IFR.

IFR Output Orders: First recalls and transitions from single- and dual-modality lists

In our experiments, we examined whether the output orders from verbal and visuo-spatial immediate memory were similar or different, and in the case of the IFR of dual-modality lists, we examined whether the output order from auditory-verbal memory and visuo-spatial memory would be independent, or whether the output orders would be highly constrained.

Consistent with Ward et al. (2010), participants in the words alone groups of our experiments tended to initiate IFR of short lists of words with the first list item, but as the list length increased, so they tended to initiate recall with one of the last few list items. Consistent with Cortis et al. (2015), participants in the dots alone groups also tended to initiate IFR of a short lists of dots with the first list item, and as the list length increased, so they tended to initiate recall with one of the last few list items. Extending the Cortis et al. findings, Experiment 1 showed that the distribution of first recalls for lists of words and dots of different length were relatively unaffected by the requirement to perform dual-modality IFR relative to the distribution of first recalls for single-modality IFR lists.

Somewhat to our surprise, there was a strong tendency to initiate recall with a word rather than a dot, and participants in the synchronized and alternating dual-modality lists of Experiments 1 and 2 tended to show highly constrained initial recalls. On trials in which participants started at the start of the list with one modality, they tended also to initiate recall from the start of the list on the other modality. Similarly, on trials in which participants started toward the end of the list with one modality, they tended also to initiate recall from the end of the list on the other modality. Indeed, throughout our experiments with synchronized and alternating lists, when we have constructed tables cross-referencing the first word recalled and the first dot recalled on a dual-modality trial, there has been clear evidence of large numbers of trials along the leading diagonal. This tendency appears to be present but much reduced in the randomized words and dots group of Experiment 3.

There was also evidence of highly constrained output orders in the analyses of transitions. In the dual-modality lists, there was clear tendency for cross-modal transitions between items from nearby serial positions (e.g., Howard & Kahana, 1999; Kahana, 1996). This was particularly the case for the synchronized words and dots group (Experiment 2), was slightly reduced in the alternating words and dots group (Experiment 2) and was still present but greatly attenuated in the randomized words and dots group (Experiment 3).

Our dual-modality findings contrast with earlier studies of IFR (for a review, see Polyn et al., 2009) that have shown that participants prefer to cluster their outputs by presentation channel (e.g., by ear, Broadbent, 1954), or by presentation modality (e.g., Murdock & Walker, 1969) or by semantic category (e.g., Bousfield, 1953). It should be noted that our studies differ from these earlier studies in that the prior studies used words or digits as stimuli, whereas we have combined auditory-verbal lists of spoken words with lists of non-verbal visual-spatial. Unlike prior experimental methodologies, it is physically possible to recall the dots and the words simultaneously, and there is no competition in our outputs to common language production mechanisms.

Our highly constrained output orders in the dual-modality groups suggest that participants are not recalling from separate visual and verbal memory systems or stores *with independent retrieval mechanisms*. If the words were encoded into distinct modality-specific memory stores (such as a phonological loop and a visuo-spatial sketchpad, Baddeley, 1986), and these slave systems had independent retrieval mechanisms then one might expect participants to output by working memory slave system, or they might prefer to recall one modality (e.g., the words) in forwards order starting at the start, but initiate recall of the other modality (e.g., the dots) from toward the end of the list. By contrast, our highly constrained output orders suggest that the auditory-verbal words and the visuo-spatial dots were encoded along a common-ordered dimension or associated with a common domain-independent timing signal , which is used at retrieval to drive forward successive outputs.

The similarities and differences between auditory-verbal IFR and visuo-spatial IFR

Throughout the three experiments, it has been possible to compare IFR of spoken words with IFR of visuo-spatial dots. Although there are some gross similarities in the output orders and serial position curves observed with words and dots, there were also some obvious and some subtle differences between IFR of the two types of stimuli. Perhaps the most obvious difference is that when a strict visuo-spatial scoring criterion was used, the participants' recall of the spoken words was far superior to the recall of the dots. There were more subtle differences between the modalities in the distribution across serial position of the first item recalled. The tendency to initiate recall with the first list item decreased more rapidly with increasing list length in the visuo-spatial lists than the auditory-verbal lists (although this can largely be attributed to an increase in errors on first outputs with increasing list length). The first recall data also suggests that when participants choose to initiate recall with one of the last four list items, there were more extended recency effects in these first recall data for the auditory-verbal stimuli than for the visuo-spatial stimuli. In addition, the extent of the primacy effects and recency effects in IFR were generally greater with words than dots.

When the visuo-spatial scoring criterion was relaxed to score as correct those responses that fell within a target radius that was twice that of the original dots (x2 criterion), we found that the visuo-spatial IFR accuracy was more closely equated to the accuracy in the auditory-verbal IFR. Importantly, although relaxing the visuo-spatial scoring criterion greatly reduced the number of errors, it did not equate the degree of primacy and recency effect, nor did it decrease the proportion of coupled first recalls, or greatly change the gross patterns of transitions at output.

Implications for theories of IFR

Current theories of IFR have rarely concerned themselves with the stimulus domain of the presented list items. Indeed, the vast majority of IFR studies over the last 50 years have examined IFR using longer lists of words as stimuli, and much of the theoretical debate in IFR and its variants has concerned the best explanations of the U-shaped serial position curve with particular emphasis on explaining recency effects, the tendency to initiate recall with one of the last few items, and the asymmetric lag recency effects in the transitions at output. Perhaps unsurprisingly given the use of word lists in IFR, there has been very little consideration in recent theorizing of the importance of the stimulus domain within these theoretical accounts of IFR. There is currently no theory of IFR that captures the growing body of visuo-spatial IFR data (e.g., Bonnani et al., 2007; Cortis et al., 2015; Dent & Smyth, 2006; Gmeindl et al., 2011), and our current understanding showing near-independence of performance limits in visuo-spatial and auditory-verbal memory suggest that theories of IFR may need to be modified if they are to adequately capture the role of stimulus representations on performance limits in IFR.

Some of our recent research examining the IFR of short lists (e.g., Grenfell-Essam & Ward, 2012; Ward et al., 2010) has already provided good reasons why an integration of the IFR and ISR literatures might be timely. As outlined earlier, participants tend to perform IFR of short lists in an “ISR-like” manner (i.e., initiating recall with the first list item and then proceeding in forwards serial order). The current dual-modality IFR data show “ISR-like” asymmetric patterns of cross-modal interference (e.g., Cowan & Morey, 2007; Morey & Mall, 2012; Morey & Miron, 2016; Morey, et al., 2013), and so our dual-modality findings now provide a further motivation to try to integrate ISR and IFR data and theory.

Perhaps the most influential account of ISR and working memory more generally, is Baddeley’s working memory model (Baddeley, 1986, 2000, 2012; Baddeley & Hitch, 1974). The model makes clear predictions that participants' immediate capacities to recall sequences of auditory-verbal words from the phonological loop might be expected to be separate from their immediate capacities to recall visuo-spatial stimuli from the visuo-spatial sketchpad. Without an Episodic Buffer (Baddeley, 2000, 2012), it would be unclear how the working memory model could account for the highly constrained output orders from cross-domain IFR. However, if both types of stimuli are additionally encoded into the Episodic Buffer, then this constraint could also be accommodated. It should be stressed that the Working Memory model has yet to be applied to IFR, that the critical Episodic Buffer remains poorly detailed, and the relationship between the Episodic Buffer, Episodic LTM and the slave systems is largely unspecified.

Our preferred explanation of this IFR data assumes that both the auditory-verbal and the visuo-spatial stimuli are associated at encoding with a common, continually-evolving, temporal context (e.g., Glenberg & Swanson, 1986; Howard & Kahana, 2002a). We assume that participants exert control over their preferred retrieval strategy at test depending upon the number of items thy are trying to recall (Tan, Ward, Paulauskaite, & Markou, 2016). Given instructions to try to recall as many list items as possible, participants attempt to initiate their recall with the first item, by reinstating the context at the start of the list context (e.g., Davelaar, et al., 2005; Henson, 1998; Metcalfe & Murdock, 1981) or via a start of list warning signal (e.g., Laming, 1999, 2010). As the length of the list increases, so successively encoded items are associated with a temporal context that is increasingly different from the start of the list context, making it harder to access the first list item. If the start of the list context cannot be retrieved, participants attempt to initiate their recall via the reinstatement of the current end of the list context (e.g., Glenberg & Swanson, 1986; Howard & Kahana, 2002a). Given that the test context is so similar to the end of list contexts, it is highly likely that this recency-based retrieval strategy is successful in first retrieving recency items. We assume that retrieval is a controlled process (e.g., Raaijmakers & Shiffrin, 1981) with participants using the same or similar retrieval cues whilst they are effective, but attempting to sample other list items via other temporal contexts when the current temporal context (or retrieved list items) are no longer effective cues. In line with Howard and Kahana (2002a), the pattern of forward-ordered transitions may reflect the possibility that contextual change at study and test may be driven by the retrieved contexts of presented and retrieved items. We believe that the highly constrained output order arises because both stimulus modalities are encoded and retrieved from a common temporal context used in the encoding and retrieval from episodic memory.

We believe that the near-independent capacities of the visuo-spatial and auditory-verbal list items may reflect fundamental differences in the encoded properties of words and dots. A number of models of ISR and more general theories of episodic memory assume that items are encoded in multiple dimensions or features (Brown et al., 2007; Cowan, 2005; Nairne, 1990; Neath, 1999; Oberauer, 2009; Oberauer & Kleigl, 2006). One could imagine that individual words and dots are represented by feature values across many dimensions. Words might possess highly diverse values across a far greater number of encoding dimensions, whilst dots might possess far less diverse values across far fewer dimensions. If words and dots share only a few overlapping dimensions, then there might be very little mutual interference at encoding and retrieval, but where interference occurs, the effect of words on dots will be greater than the effect of dots on words.

Finally, our data replicate the findings of Cortis et al. by showing greater primacy and recency effects with the IFR of words relative to dots. In addition, our data show a more graded recency effect in the probability of first recall data for words than dots. Although these differences may be explained by the greater diversity of words relative to dots, these differences might also reflect the greater sensitivity of auditory-verbal stimuli in immediate recall tasks to sequential ordering and temporal grouping (e.g., Farrell, 2012; Macken, Taylor, Kozlov, Hughes, & Jones, 2016; Spurgeon et al., 2015) and rehearsal (e.g., Barrouillet & Camos, 2015; Tan & Ward, 2000; 2008), whereas the IFR of visuo-spatial dots may be more sensitive to spatial proximity (cf. Bor et al., 2003; Gmeindl et al., 2011).

Summary and Conclusions

In summary, we have produced novel and theoretically interesting data sets examining the IFR of single- and dual-modality lists of auditory-verbal and /or visuo-spatial stimuli. We find (1) near-independent IFR performance limitations, with only a relative small dual-task decrement in the recall of visuo-spatial but not auditory-verbal stimuli and (2) highly constrained output orders in the IFR of the two modalities. We also find (3) many similarities and a few differences in the patterns of output associated with IFR of verbal and visuo-spatial stimuli. We believe that it is increasingly timely to integrate theories of IFR with theories of ISR (and working memory). We believe that our findings suggest that verbal and visuo-spatial stimuli are associated with a common continuously-evolving temporal context, but the representations of verbal and visuo-spatial stimuli are sufficiently dissimilar to result in near-independent performance limitations.

References

Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2009). Cross-modal binding and working memory. *Visual Cognit*ion, *17*, 83-102.

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

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

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory*? Trends in Cognitive Sciences, 4,* 417 – 423.

Baddeley, A. D. (2012). Working memory: Theories, models and controversies. *Annual Review of Psychology, 63*, 1-29.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation, 8*, 47-89.

Baddeley, A.D., & Logie, R.H. (1999). The multicomponent model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28-61). Cambridge: Cambridge University Press.

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction.* Hove, UK: Psychology Press.

Bonanni, R., Pasqualetti, P., Caltagirone, C., & Carlesimo, G. A. (2007). Primacy and recency effects in immediate free recall of sequences of spatial positions. *Perceptual and motor skills*, *105*(2), 483-500.

Bor, D., Duncan, J., Wiseman, R. J., & Owen, A. M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron*, *37*, 361-367.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49*, 229–240.

Broadbent, D. E. (1954). The role of auditory localization in attenton and memory span. *Journal of Experimental Psychology, 47*, 191-196.

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

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and non-verbal stimuli. *Journal of Experimental Psychology: Learning, Memory and Cognition, 41,* 1179-1214.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin*, *104*, 163-191.

Cowan, N. (1998). *Attention and memory: An integrated framework.* Oxford University Press.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, *24*, 154-176.

Cowan, N. (2005). *Working memory capacity*. East Sussex: Psychology Press.

Cowan, N., & Morey, C. C. (2007). How can dual-task working memory retention limits be investigated? *Psychological Science, 18*, 686-688.

Crowder, R. G., &: Morton, J. (1969). Precategorical Acoustic Storage (PAS). *Perception & Psychophysics, 5*, 365-373.

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

Dent, K., & Smyth, M.M. (2006). Capacity limitations and representational shifts in spatial short-term memory. *Visual Cognition, 13,* 529-572.

Depoorter, A., & Vandierendonck, A. (2009). Evidence for modality-independent order coding in working memory. *The Quarterly Journal of Experimental Psychology, 62*, 531-549.

Doherty, J. M., & Logie, R. H. (2016). Resource-sharing in multiple-component working memory. *Memory & Cognition*, *44*, 1157-1167.

Dornbush, R. L. (1968). Input variables in bisensory memory. *Perception and Psychophysics, 4*, 41-44.

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

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

Gibson, B. S., Gondoli, D. M., Flies, A. C., Dobrzenski, B. A., & Unsworth, N. (2009). Application of the dual-component model of working memory to ADHD. *Child Neuropsychology*, *16*(1), 60-79.

Gibson, B. S., Gondoli, D. M., Flies, A. C., Dobrzenski, B. A., & Unsworth, N. (2009). Application of the dual-component model of working memory to ADHD. *Child Neuropsychology*, *16*, 60-79.

Gibson, B. S., Gondoli, D. M., Johnson, A. C., & Robison, M. K. (2014). Recall initiation strategies must be controlled in training studies that use immediate free recall tasks to measure the components of working memory capacity across time. *Child Neuropsychology*, *20*, 539-556.

Gibson, B. S., Gondoli, D. M., Johnson, A. C., Steeger, C. M., Dobrzenski, B. A., & Morrissey, R. A. (2011). Component analysis of verbal versus spatial working memory training in adolescents with ADHD: A randomized, controlled trial. *Child Neuropsychology*, *17*, 546-563.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, *5*, 351-360.

Glenberg, A. M., &: Swanson, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 12*, 3-15.

Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & cognition, 39*, 37-46.

Greene, R. L. (1989). Immediate serial recall of mixed-modality lists. *Journal of Experimental Psychology: Learning, Memory & Cognition, 15*, 266-274.

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: The role of list length, strategy use, and test expectancy. *Journal of Memory and Language, 67,* 106-148.

Grenfell-Essam, R. & Ward, G. (2015). The effect of selective attention and a stimulus prefix on the output order of immediate free recall of short and long lists. *Canadian Journal of Experimental Psychology, 69(1)*, 1-16.

Grenfell-Essam, R., Ward, G., & Tan, L. (2013). The role of rehearsal on the output order of immediate free recall of short and long lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 317-347.

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 556-569.

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies?. *Journal of Experimental Psychology: General*, *143*, 575.

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

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

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494-500

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2*5, 1-19.

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

Howard, M. W., & Kahana, M. J. (2002b). When does semantic similarity help episodic retrieval?. *Journal of Memory and Language*, *46*(1), 85-98.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140*, 339-373.

Jahnke, J. C. (1965). Primacy and recency effects in serial-position curves of immediate recall. *Journal of Experimental Psychology*, *70*, 130-132.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a gestural view of short-term memory. *Journal of Memory and Language, 54*, 265-281.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological, and is it a store? *Journal of Experimental Psychology: Learning, Memory & Cognition, 30*, 656-674.

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

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

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

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

Laming, D. (2010). Serial position curves in free recall. *Psychological Review, 117*, 93-113.

Logie, R. H. (1995). *Visuo-spatial Working Memory.* Hove, East Sussex, UK: Lawrence Erlbaum Associates.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science, 20,* 240-245.

Macken, B., Taylor, J. C., Kozlov, M. D., Hughes, R. W., & Jones, D. M. (2016). Memory as embodiment: the case of modality and serial short-term memory. *Cognition*, *155*, 113-124.

Margrain, S. A. (1967). Short-term memory as a function of input modality. *The Quarterly Journal of Experimental Psychology*, *19*, 109-114.

Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *20*(2), 161-189.

Morey, C. C., & Mall, J. T. (2012). Cross-domain interference costs during concurrent verbal and spatial serial memory tasks are asymmetric. *The Quarterly Journal of Experimental Psychology, 65*, 1777-1797.

Morey, C. C., & Miron, M. D. (2016). Spatial sequences, but not verbal sequences, are vulnerable to general interference during retention in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1907.

Morey, C. C., Morey, R. D., van der Reijden, M., & Holweg, M. (2013). Asymmetric cross-domain interference between two working memory tasks: Implications for models of working memory. *Journal of Memory and Language, 69*, 324-348.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482-488.

Murdock, B. B., & Carey, S. T. (1972). Release from interference in single-trial free recall. *Journal of Verbal Learning and Verbal Behaviour, 11*, 398-402.

Murdock, B. B., & Walker, K. D. (1969). Modality effects in free recall *Journal of Verbal Learning and Verbal Behaviour, 8*, 665-676. Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology, 21*, 263-276.

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

Neath, I. (1999). Modeling the disruptive effects of irrelevant speech on order information. *International Journal of Psychology, 34*, 410-418.

Nilsson, L. G., Ohlsson, K., & Rönnberg, J. (1977). Capacity differences in processing and storage of auditory verbal and visual input. In S. Dornic (Ed.), *Attention and Performance* (Vol. VI, pp. 629-645). Hillsdale, NJ: Erlbaum.

Oberauer, K., (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory and Cognition, 37,* 346-357.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language, 55*, 601-626.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, and Winston.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45,* 255–287.

Parmentier, F. B. R. (2011). Exploring the determinants of memory for spatial sequences. In A. Vandierendonck & A. Szmalec (Eds.), *Spatial working memory* (pp. 67-86). New York: Psychology Press.

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

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall, *Quarterly Journal of Experimental Psychology, 17,* 132 - 138.

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

Roberts, W. A. (1972). Free recall of word lists varying in length and rate of presentation: A test of total-time hypotheses. *Journal of Experimental Psychology, 92,* 365-372.

Sanders, A. F., & Schroots, J. J. F. (1969). Cognitive categories and memory span. III. Effects of similarity on recall. *The Quarterly Journal of Experimental Psychology*, *21*, 21-28.

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

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*, 1-29.

Spurgeon, J., Ward, G., & Matthews, W. J. (2014a). Examining the relationship between immediate serial recall and immediate free recall: Common effects of Phonological Loop variables, but only limited evidence for the Phonological Loop. *Journal of Experimental Psychology: Learning, Memory & Cognition, 40*, 1110-1141.

Spurgeon, J., Ward, G., & Matthews, W.J. (2014b). Why do participants initiate free recall of short lists with the first list item? Toward a general episodic memory explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40, 1551-1567.*

Spurgeon, J., Ward, G., Matthews, W.J., & Farrell, S. (2015). Can the effects of temporal grouping explain the similarities and differences between free recall and serial recall? *Memory & Cognition, 43*, 469-488.

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

Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review, 15*, 535-542.

Tan, L., Ward, G., Paulauskaite, L., Markou, M. (2016). Beginning at the beginning: Recall order and the number of words to be recalled. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42, 1282-1292*

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104-132.

Vandierendonck, A. (2016). Modality independence of order coding in working memory: Evidence from cross-modal order interference at recall. *The Quarterly Journal of Experimental Psychology, 69*, 161-179.

Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory and Cognition, 30*, 885-892.

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory, 13*, 308-317.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the Relationship between Free Recall and Immediate Serial Recall: The Effects of List Length and Output Order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1207-1241.

Waugh, N. C. (1967). Presentation time and free recall. *Journal of Experimental Psychology*, *73*, 39-44.

Table 1
The Probability of First Recall data from Experiment 1. The frequencies with which participants initiated recall with items presented at different input serial positions as a function of group, type of stimulus, and list length (LL). Frequencies of initial outputs for the first and last four serial positions are in **bold** font.

	Serial Position																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	V/E
Words alone group																	
LL1	**196**																4
LL2	**193**	**3**															4
LL4	**174**	**9**	**5**	**5**													7
LL8	**85**	7	11	8	**11**	**14**	**21**	**37**									6
LL16	**51**	5	4	1	4	1	3	4	5	3	11	6	**9**	**15**	**30**	**42**	6
Words in the Synchronized words and dots group																	
LL1	**198**																2
LL2	**192**	**2**															6
LL4	**171**	**5**	**4**	**13**													7
LL8	**63**	6	6	6	**17**	**18**	**17**	**61**									6
LL16	**22**	2	2	3	3	3	2	5	5		6	5	**8**	**12**	**24**	**90**	8
Dots alone group																	
LL1	**170**																30
LL2	**129**	**18**															53
LL4	**95**	**10**	**4**	**5**													86
LL8	**55**	7	2	5	**1**	**3**	**5**	**32**									90
LL16	**19**	4	3				1	4	3	1	4	4	**4**	**6**	**6**	**58**	83
Dots in the Synchronized words and dots group																	
LL1	**171**																29
LL2	**135**	**7**															58
LL4	**92**	**3**	**3**	**13**													89
LL8	**37**	2	2	6	**6**	**7**	**8**	**49**									83
LL16	**11**	1	4	1	2	3		6	3	2	1	2	**3**	**7**	**14**	**62**	78

Note: V/E refers to Voids / Errors. Voids refer to trials in which there were no responses, and so no first item or where the first recall was incorrect (an Error)

Table 2

Detailed analyses of the Probability of First Recall data for the synchronized words and dots group from Experiment 1. The frequencies with which participants initiated recall with words and dots from different input serial positions for each list length. The rows represent the serial positions of the first words recalled and the columns represent the serial positions of the first dots recalled. Values in **bold** font represent trials with highly-coupled initial recalls in which the serial position of the first word recalled matched the serial position of the first dot recalled.

	visuo-spatial stimuli	
verbal stimuli		
List Length 1	1	V/E
1	**169**	29
V/E	2	

List Length 2	1	2	V/E
1	**130**	4	58
2		**2**	
V/E	5	1	

List Length 4	1	2	3	4	V/E
1	**83**	1	2	1	84
2	3				2
3		1	**1**		2
4				**12**	1
V/E	6	1			

List Length 8	1	2	3	4	5	6	7	8	V/E
1	**31**					2	1		29
2		**1**							5
3			**2**	2					2
4	1			**2**		1			2
5	2	1			**6**		2	1	5
6						**4**	1	1	12
7	1			1			**4**	2	8
8	1							**43**	17
V/E	1			1				2	2

List Length 16	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	V/E
1	**7**		1														12
2			1														1
3	1					1											
4			1	**1**													1
5	1									1							1
6						**2**			1								
7																	2
8					1			**3**									1
9															1		4
10																	
11									1		**1**			1			3
12												**1**	1				3
13										1			**2**				5
14														**2**	3		7
15					1			1	1					2	**9**	2	8
16	1	1	1					1				1		2		**58**	25
V/E	1							1								1	5

Table 3

Experiment 2: The frequency of trials where recall was initiated with words from each serial position, where the dual-modality groups were separated by modality

	Serial Position						
	1	2	3	4	5	6	Void/Error
Words only	425	62	57	68	142	225	21
Words Synchronized	373	45	37	34	94	397	20
Words Alternating	589	33	39	39	102	160	38
Dots only	245	38	33	24	48	197	415
Dots Synchronized	180	21	16	19	58	318	388
Dots Alternating	255	30	27	22	39	108	519

Table 4

Experiment 2: The total number of the first within-modality responses for the recalled visuo-spatial dots and words as a function of serial position for the synchronized onsets and alternating words and dots groups respectively.

	Visuo-spatial stimuli						
Verbal Stimuli	1	2	3	4	5	6	Void/Error
Synchronized Onsets Group							
1	153	5	3	3	2	8	199
2	4	12	1	2	2	1	23
3	2		10	1	2	3	19
4	3		1	9	3	2	16
5	4	2		1	42	12	33
6	9	2	1	2	5	289	89
Error/Void	5			1	2	3	9
Alternating Words and Dots Group							
1	219	14	14	6	4	9	323
2	2	8	1	1			21
3	5	5	8				21
4	3	1	2	10	3		20
5	7		1	3	29	14	48
6	8	1	1	2	3	83	62
Error/Void	11	1				2	24

Table 5
Experiment 3: The total number of trials that were started at each of the serial position for each group and stimulus modality.

	Serial Position						
	1	2	3	4	5	6	Void/Error
Words only	514	40	41	69	112	180	44
Word Randomized	570	37	41	49	99	169	35
Dots only	296	24	18	22	31	99	510
Dots - Randomized	237	32	26	30	43	109	523

Table 6
Experiment 3: The total number of the first within-modality responses for the recalled visuo-spatial dots and words as a function of serial position for the random presentation group.

	Visuo-spatial stimuli						
Verbal Stimuli	1	2	3	4	5	6	Void/Error
1	174	17	18	10	15	36	301
2	10	1	1	4	2	8	12
3	9	1		4	2	3	22
4	8	3		4	2	7	25
5	14	1	3	2	9	19	51
6	14	5	5	5	13	32	96
Error/Void	8	4		1		4	16

Figure Captions

Figure 1. Overall accuracy data from Experiment 1. The mean proportion of items recalled at each list length for the words recalled in the words only group, for the dots recalled in the Dots alone group, and for the words and the dots from the synchronized words and dots group.

Figure 2. Serial position curves for each list length from Experiment 1. Panel 2A shows the auditory-verbal IFR data from the Words alone group and the words from the synchronized words and dots group. Panel 2B shows the visuo-spatial IFR data from the Dots alone group and the dots from the synchronized words and dots group.

Figure 3. The probability of first recall (PFR) data from Experiment 1. For the auditory-verbal list items in the Words alone and Words from the Synchronized words and dots groups, this is the proportion of trials in which the first spoken word was the word presented in serial position 1. For the visuo-spatial list items in the Dots alone and Dots from the Synchronized words and dots groups, this is the proportion of trials in which the first mouse click response was the dot presented in serial position 1.

Figure 4. The frequency with which different lag transitions were made in the Synchronized words and dots group for Experiment 1. Transitions could be made within-modalities (unfilled plot symbols) or cross-modalities (filled plot symbols). The Lag refers to the difference between the within-modality serial position of successive items (obtained by subtracting the serial position of output *n* from the serial position of output *n+1)* such that smaller lags denote transitions between items from neighboring serial positions in the experimenter's list. Thus, a Lag of 0 refers to transitions between dots and words or between words and dots that were presented at the same time, and a Lag of +1 refers to forward-ordered transitions between stimuli from consecutive serial positions.

Figure 5. Overall accuracy data from Experiment 2 (Figure 5A) and Experiment 3 (Figure 5B), segregated by stimulus modality and group. In Figure 5A, the triangles plot the mean proportion of words correctly recalled from the words alone group, from the Synchronized words and dots group and from the Alternating words and dots group. The circles plot the mean proportion of dots correctly recalled from the dots alone group, from the Synchronized words and dots group and from the Alternating words and dots group. In Figure 5B, the triangles plot the mean proportion of words correctly recalled from the words alone group and from the Randomized words and dots group. The circles plot the mean proportion of dots correctly recalled from the dots alone group and from Randomized words and dots group.

Figure 6. Serial position curves for each stimulus modality and each group from Experiment 2. Panel 6A shows the auditory-verbal IFR data from the Words alone group, the words from the synchronized words and dots group, and the words from the Alternating words and dots group. Panel 6B shows the visuo-spatial IFR data from the Dots alone group, the dots from the synchronized words and dots group, and the dots from the Alternating words and dots group.

Figure 7. The Lag-Conditionalized Response Probabilities (Lag-CRP) plots from the dual-modality groups in Experiment 2. Panel 7A shows the transition data from the Synchronized words and dots group; Panel 7B shows the transition data from the Alternating words and dots group. Transitions could be made within-modalities (unfilled plot symbols) or cross-modalities (filled plot symbols). The Lag refers to the difference between the within-modality serial position of successive items (obtained by subtracting the serial position of output *n* from the serial position of output *n+1)* such that smaller lags denote transitions between items from neighboring serial positions in the experimenter's list. A Lag of +1 refers to forward-ordered transitions between stimuli from consecutive serial positions. For the Synchronized words and dots group, there were 6 serial positions and so the Lag could vary from between -5 and +5. Note that a Lag of 0 refers to transitions between dots and words or between words and dots that were presented at the same time.

For the Alternating words and dots group, there were effectively 12 serial positions and so the Lag could vary from between -11 and +11. Note that it was not possible for a Lag of 0 as only one stimulus was ever presented at a given serial position.

Figure 8. The serial position curves for each group and stimulus modality in Experiment 3. The data from the randomized words and dots group was segregated by verbal and visuo-spatial modality and compared to its respective single-modality group performance.

Figure 9. The Lag-Conditionalized Response Probabilities (Lag-CRP) plots from the Randomized words and dots group in Experiment 3. Transitions could be made within-modalities (unfilled plot symbols) or cross-modalities (filled plot symbols). The Lag refers to the difference between the within-modality serial position of successive items (obtained by subtracting the serial position of output *n* from the serial position of output *n+1)* such that smaller lags denote transitions between items from neighboring serial positions in the experimenter's list. There were effectively 12 different serial positions, so the lag can vary from between -11 and +11.

Figure 10. The effects of relaxing the strict accuracy criterion used to score the recall of the dot stimuli in Experiments 2 and 3. Using the strict criterion of x1, a dot is considered to be correctly recalled only if a mouse-click falls within its circumference (within x1 radius of the dot's center). This criterion can be relaxed to score near-misses as correct, such that a dot is considered to be correctly recalled if a mouse-click falls between x1 and x2 radii of the dot's center). Figures 10A and 10B show how the frequency of total errors decrease as the scoring criterion is relaxed in the groups yielding visuo-spatial data in Experiments 2 and 3, respectively. Figures 10C and 10D show how the accuracy increases as the scoring criterion is relaxed in the groups yielding visuo-spatial data in Experiments 2 and 3, respectively. Finally, Figures 10E and 10F show how the proportion of tightly-coupled first word and first dot responses vary as the scoring criterion is relaxed in the groups yielding dual-modality data in Experiments 2 and 3, respectively. A tightly-coupled response is

where the first word recalled is from the same within-modality serial position as the first dot recalled.

Figure 11. The serial position curves using the more lenient scoring criterion (x2) for the visuo-spatial dot stimuli for each group in Experiment 2 (Panel 11A) and in Experiment 3 (Panel 11B). Using the x2 scoring criterion, a mouse-click is considered correct if it falls within a target radius that is twice that of the stimulus dot.

Figure 12. The Conditionalized Response Probabilities as a function of lag transitions for the dual-modality groups of Experiments 2 and 3 using the more lenient criterion (x2) for scoring the visuo-spatial responses. Using the x2 scoring criterion, a mouse-click is considered correct if it falls within a target radius that is twice that of the stimulus dot. Panel 12A shows the Lag-CRP plots for the Synchronized words and dots group of Experiment 2, Panel 12B shows the Lag-CRP plots for the Alternating words and dots group of Experiment 2, and Panel 12C shows the Lag-CRP plots for the Randomized words and dots group of Experiment 3.

Figure 1

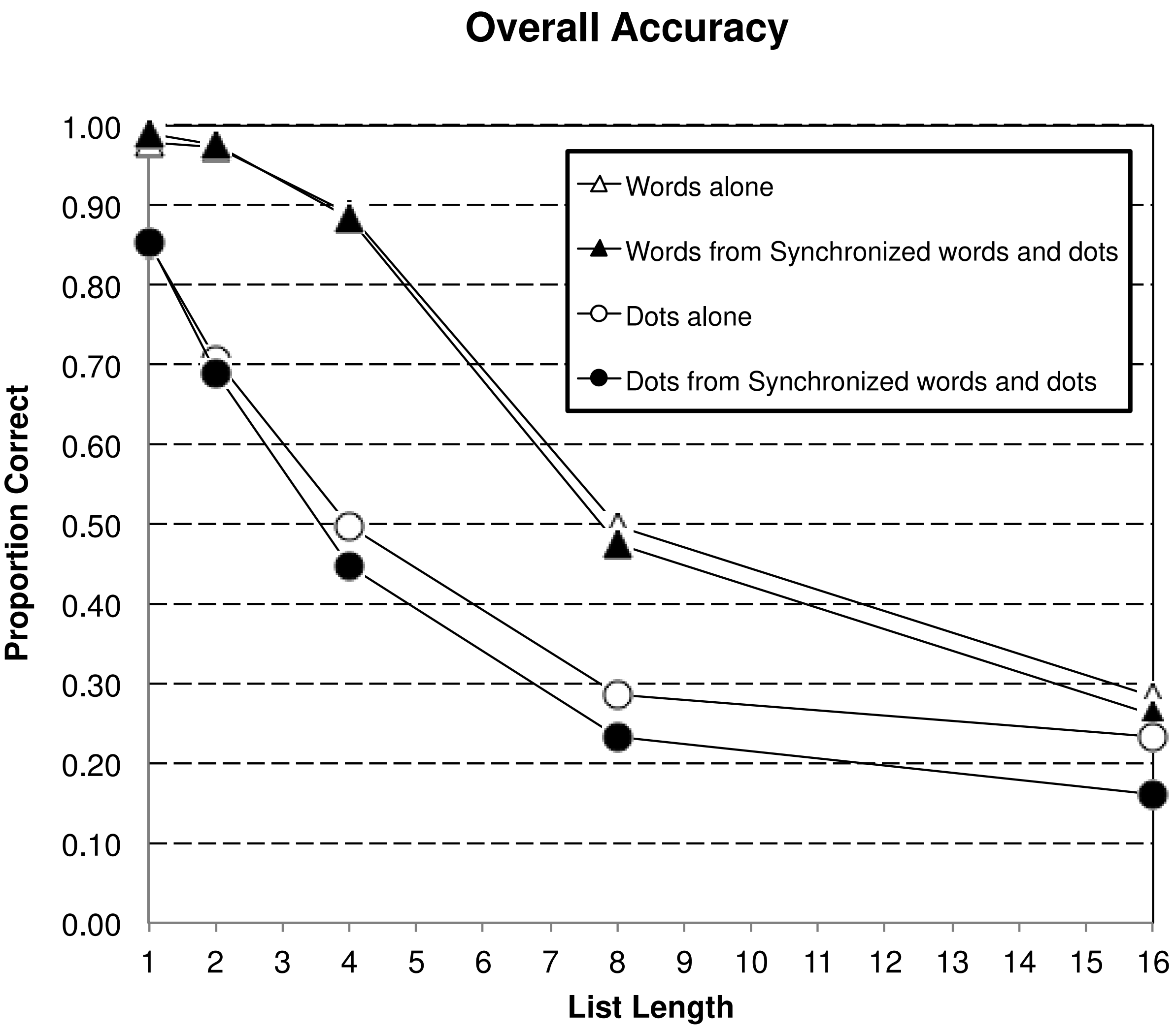
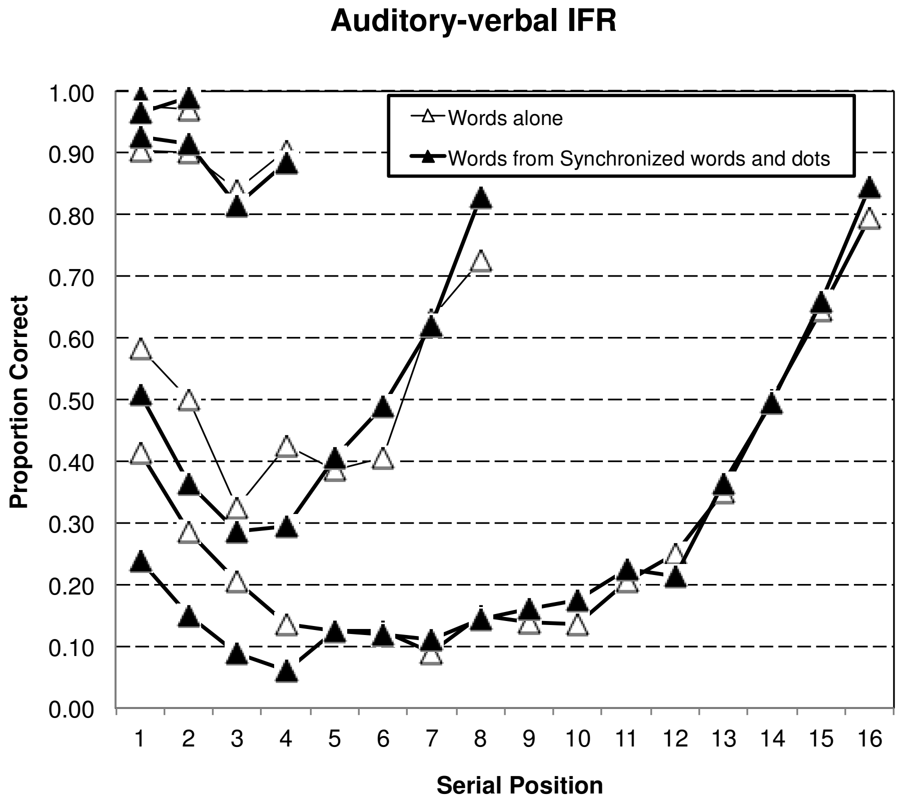

Overall Accuracy
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
List Length
Words alone
Words from Synchronized words and dots
Dots alone
Dots from Synchronized words and dots

Figure 2


Auditory-verbal IFR
Words alone
Words from Synchronized words and dots
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Serial Position

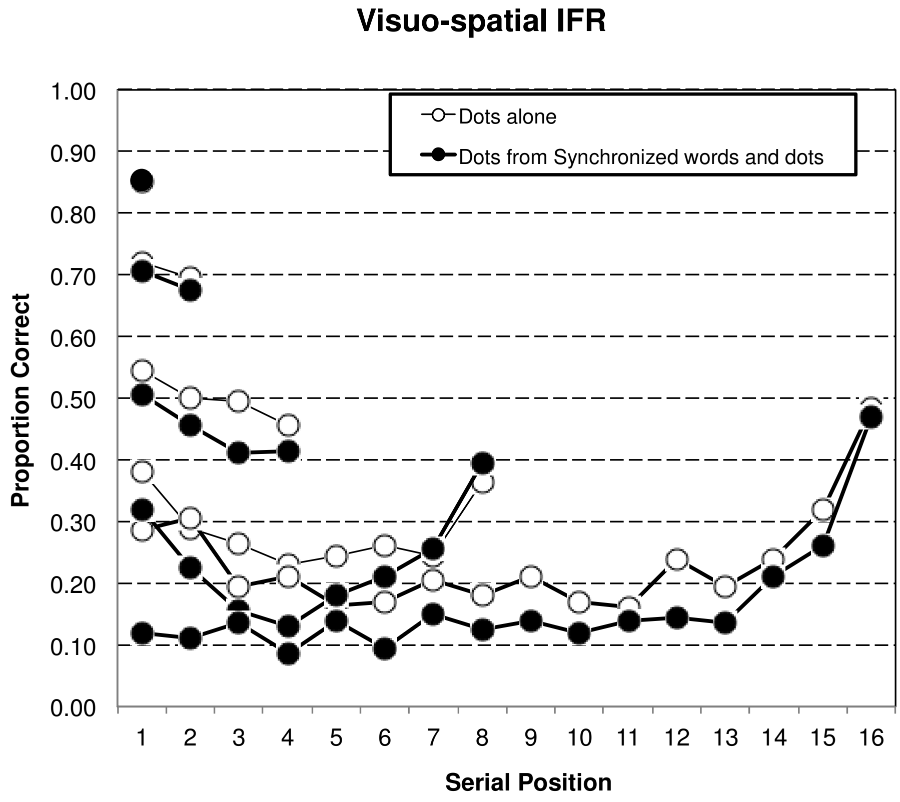

Visuo-spatial IFR
Dots alone
Dots from Synchronized words and dots
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Serial Position

Figure 3

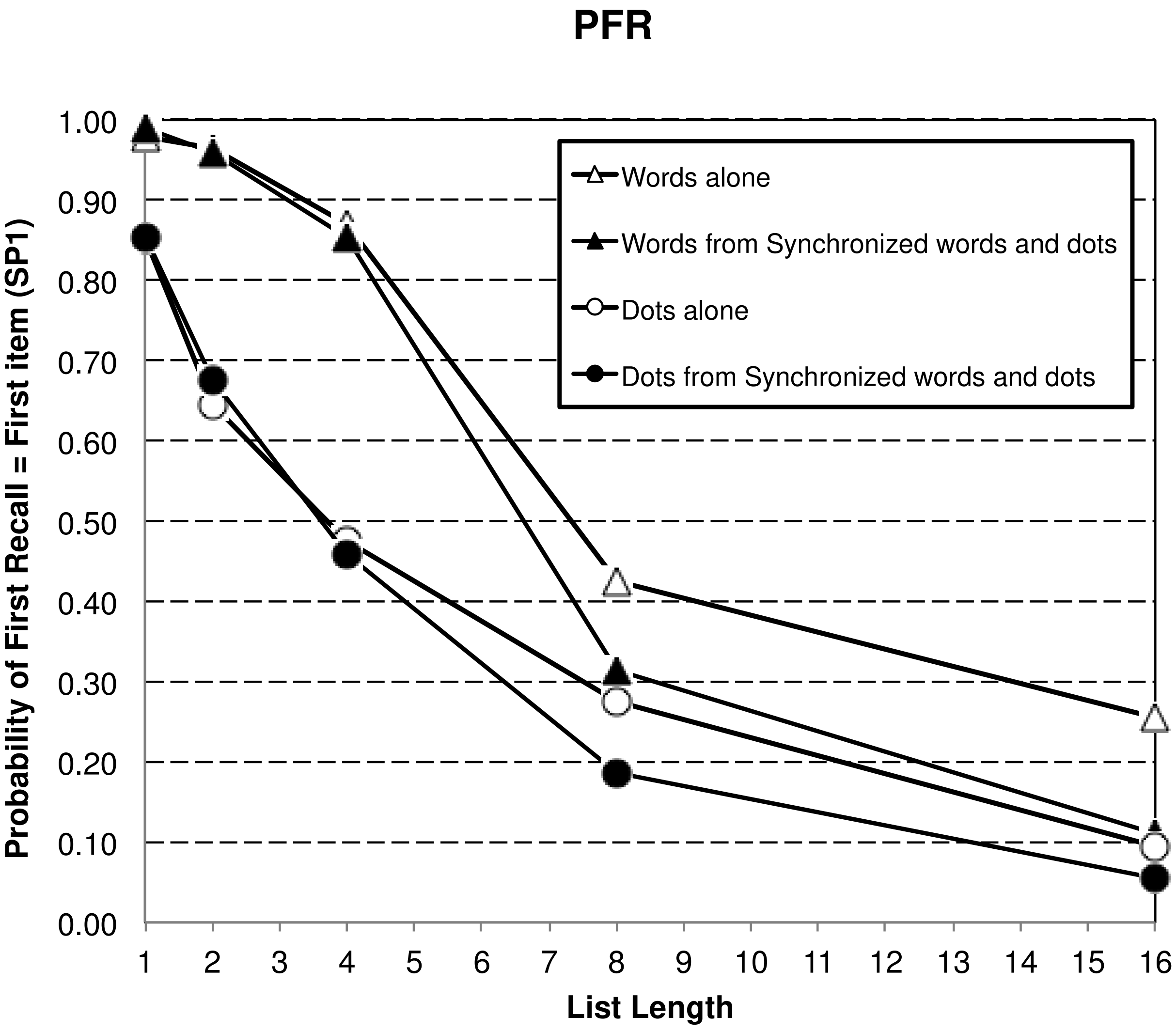

PFR
Probability of First Recall = First item (SP1)
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
List Length
Words alone
Words from Synchronized words and dots
Dots alone
Dots from Synchronized words and dots

Figure 4

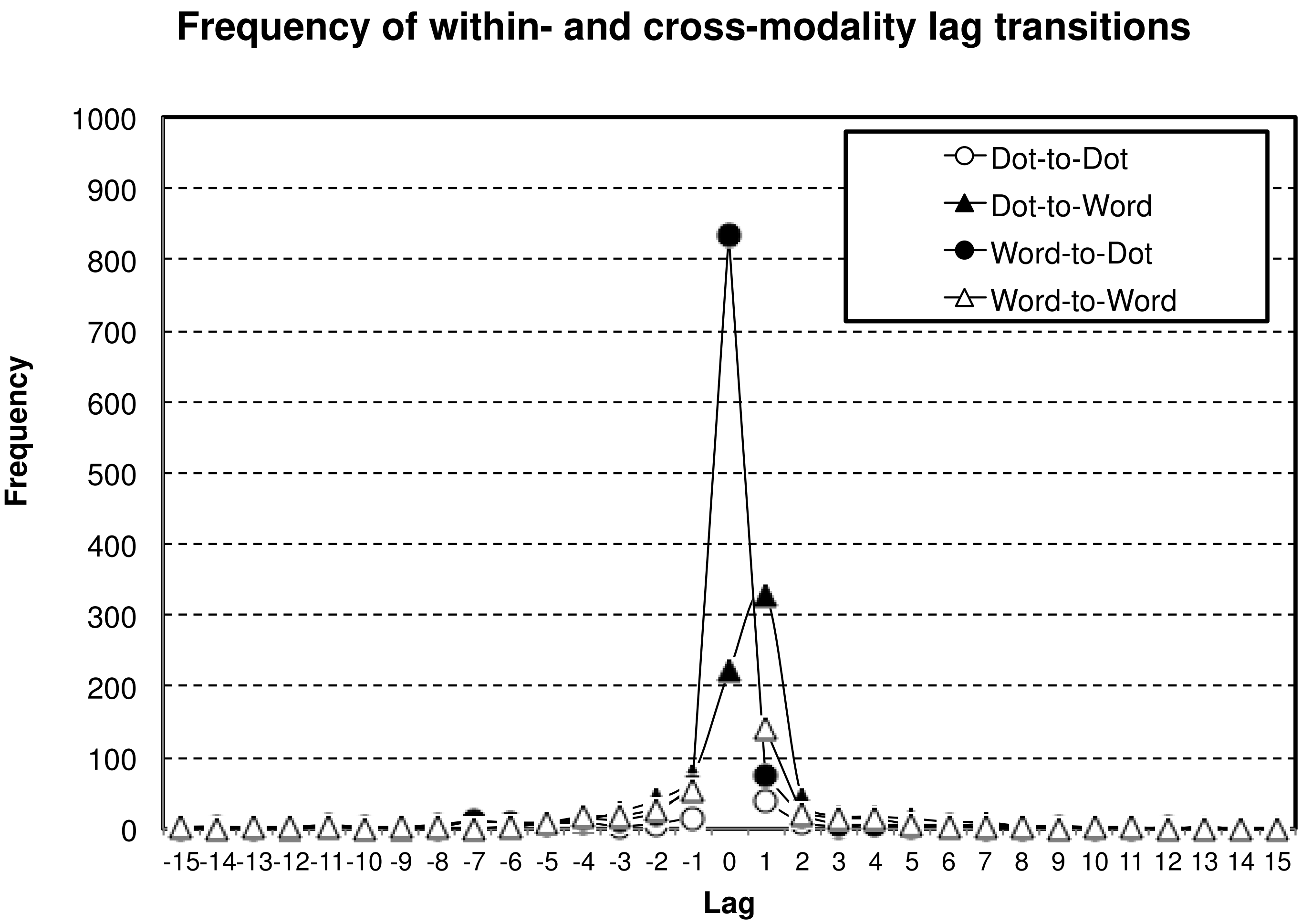

Frequency of within- and cross-modality lag transitions
Frequency
1000
900
800
700
600
500
400
300
200
100
0
-15
-14
-13
-12
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Lag
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word

Figure 5

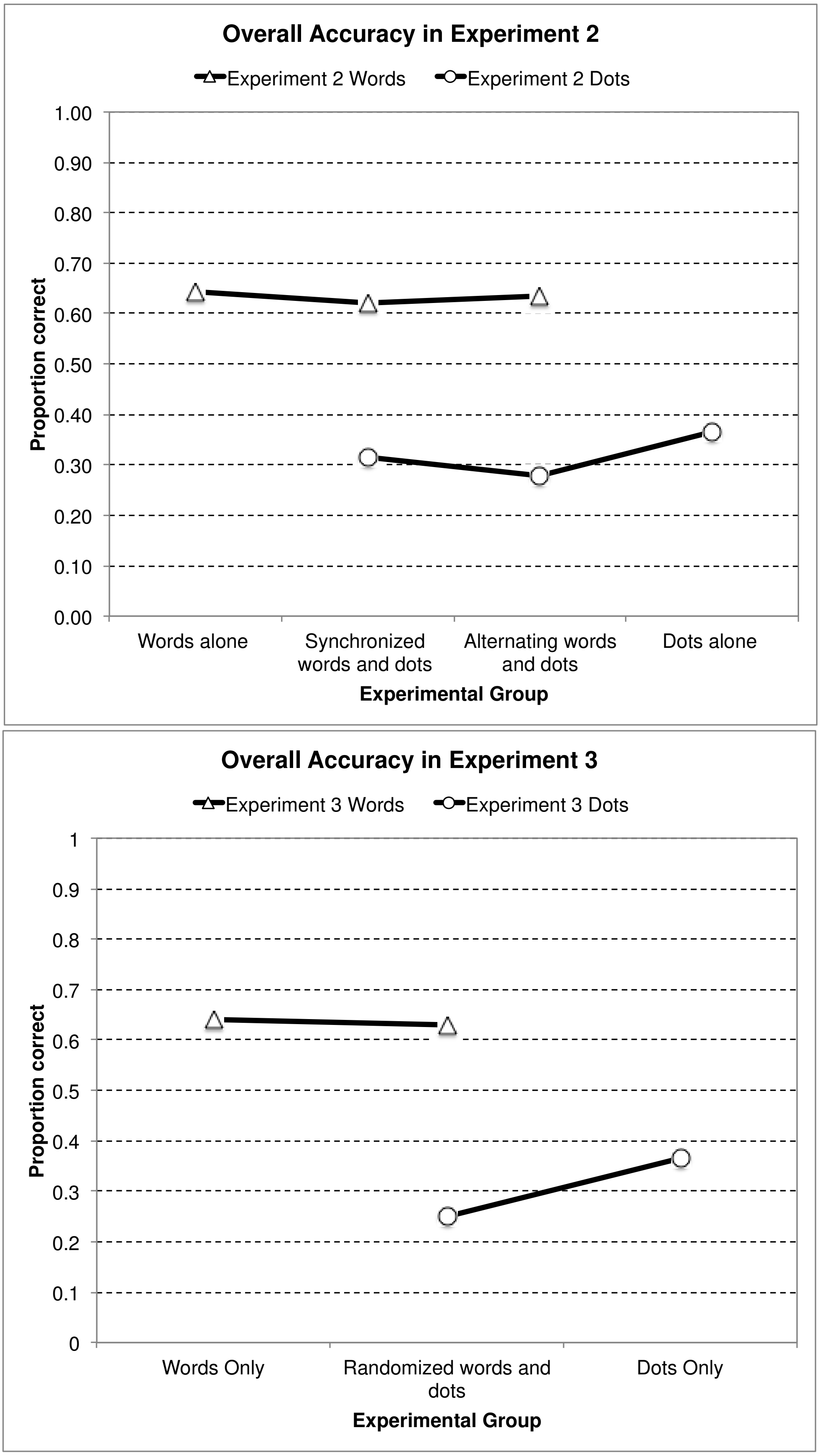
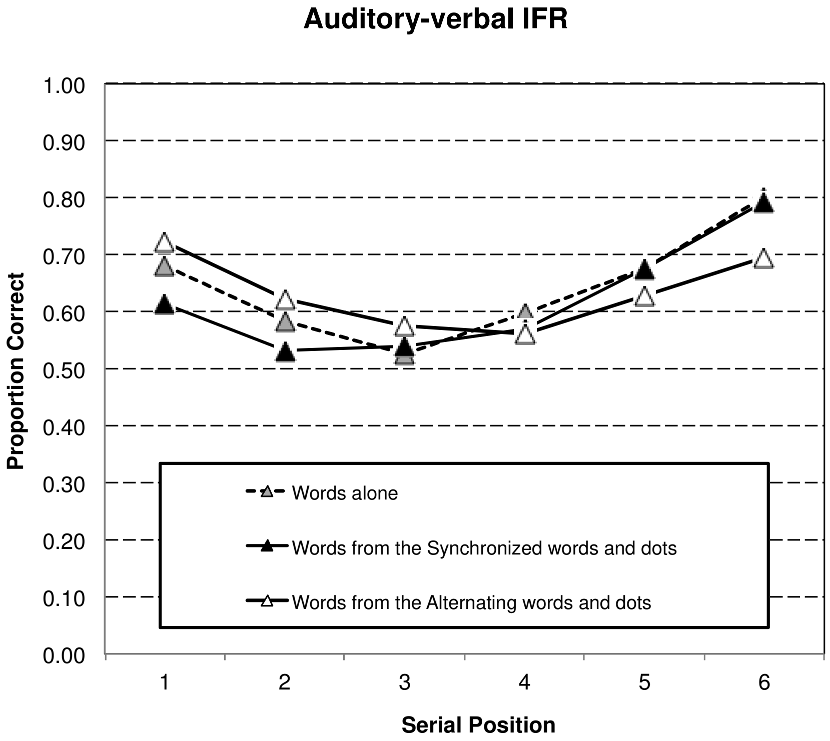

Overall Accuracy in Experiment 2
Experiment 2 Words
Experiment 2 Dots
Proportion correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Words alone
Synchronized words and dots
Alternating words and dots
Dots alone
Experimental Group
Overall Accuracy in Experiment 3
Experiment 3 Words
Experiment 3 Dots
Proportion correct
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Words Only
Randomized words and dots
Dots Only
Experimental Group

Figure 6


Auditory-verbal IFR
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Serial Position
1
2
3
4
5
6
Words alone
Words from the Synchronized words and dots
Words from the Alternating words and dots

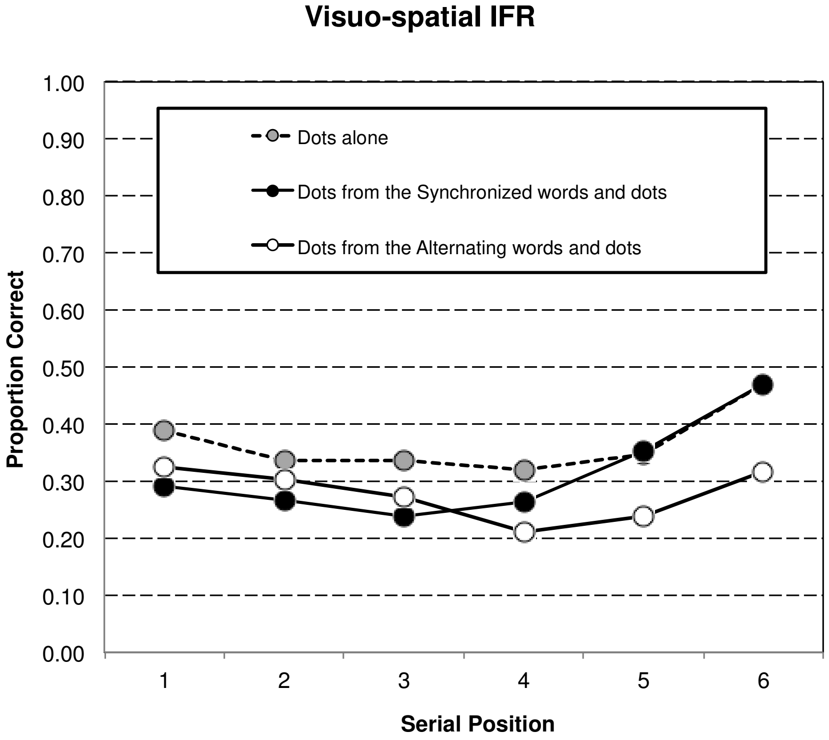

Visuo-spatial IFR
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Serial Position
1
2
3
4
5
6
Dots alone
Dots from the Synchronized words and dots
Dots from the Alternating words and dots

Figure 7

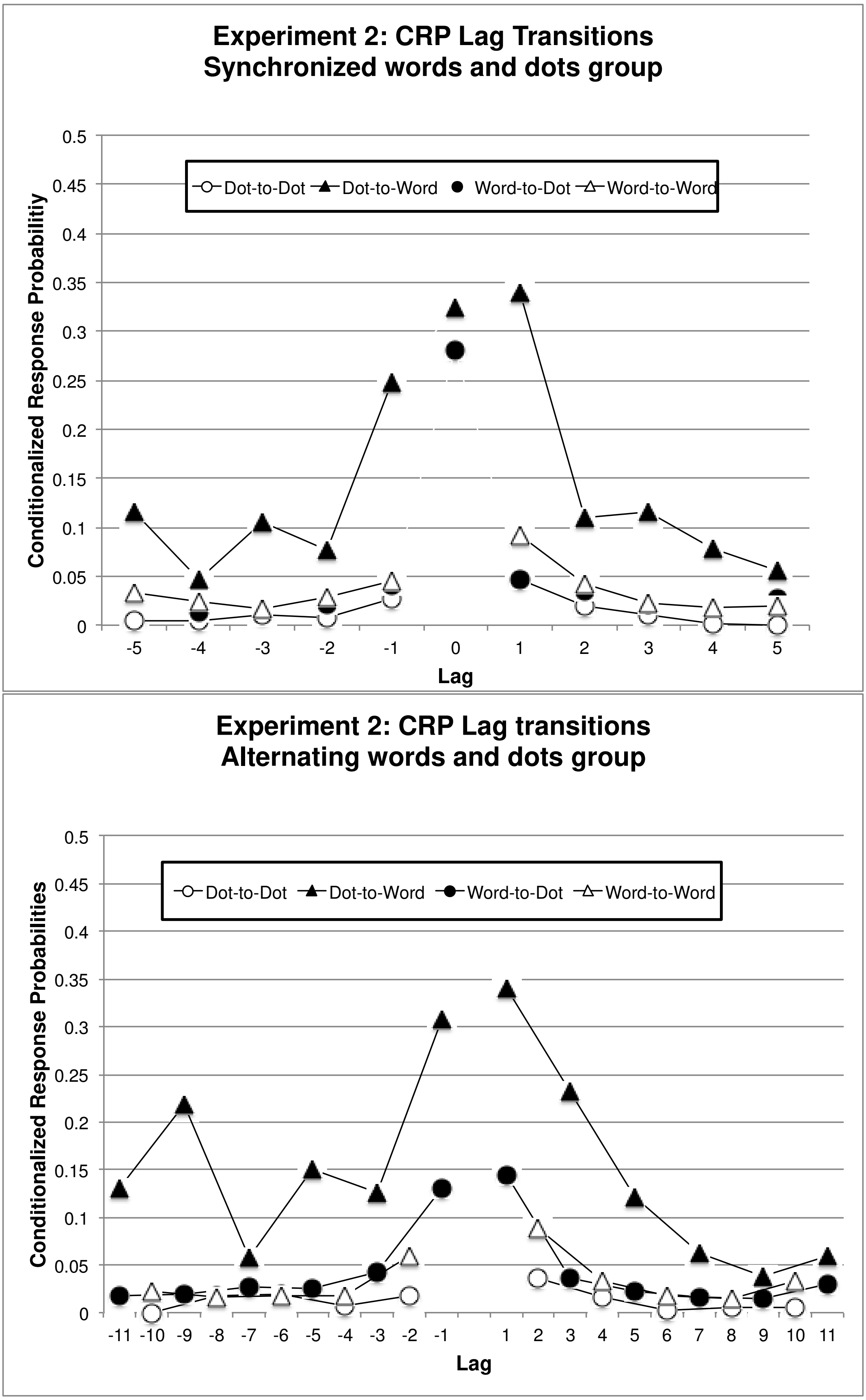

Experiment 2: CRP Lag Transitions
Synchronized words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilitiy
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5
-4
-3
-2
-1
0
1
2
3
4
5
Lag
Experiment 2: CRP Lag transitions
Alternating words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilities
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10
11
Lag

Figure 8

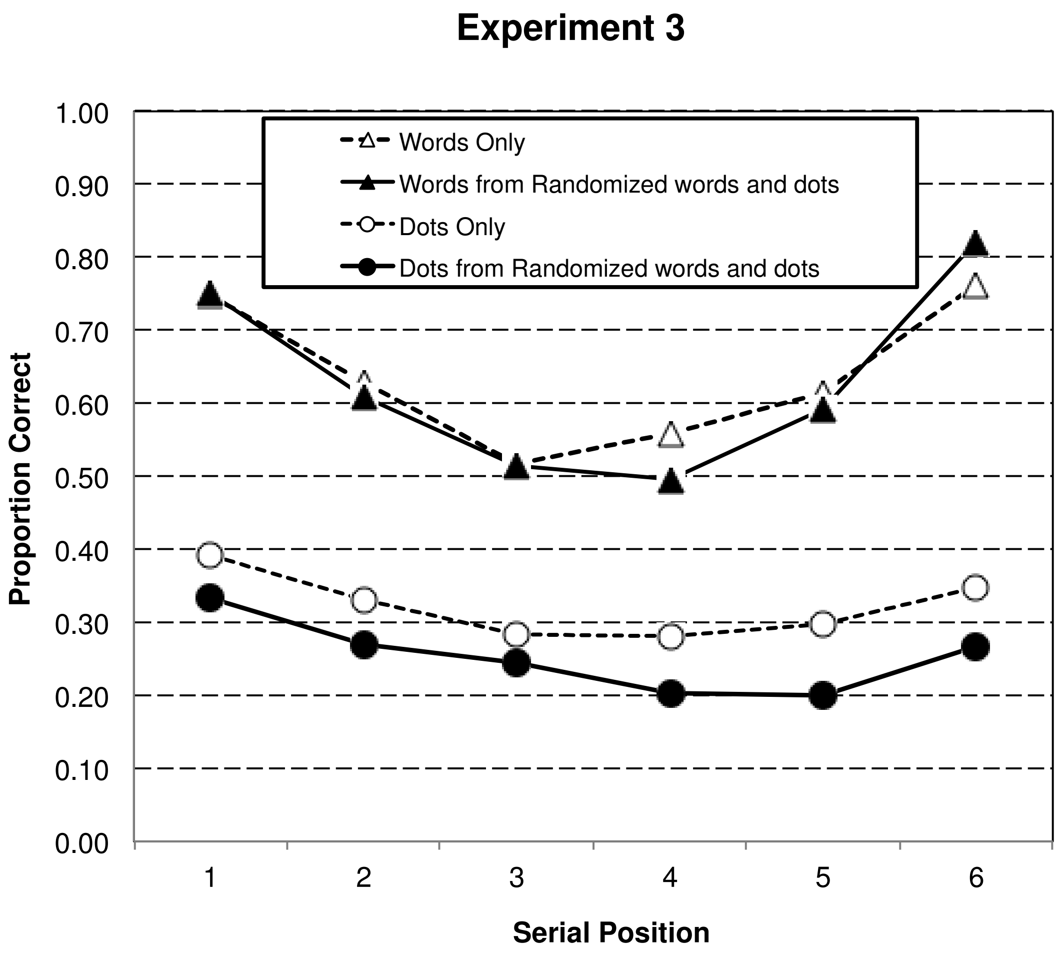

Experiment 3
Words Only
Words from Randomized words and dots
Dots Only
Dots from Randomized words and dots
Proportion Correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
1
2
3
4
5
6
Serial Position

Figure 9

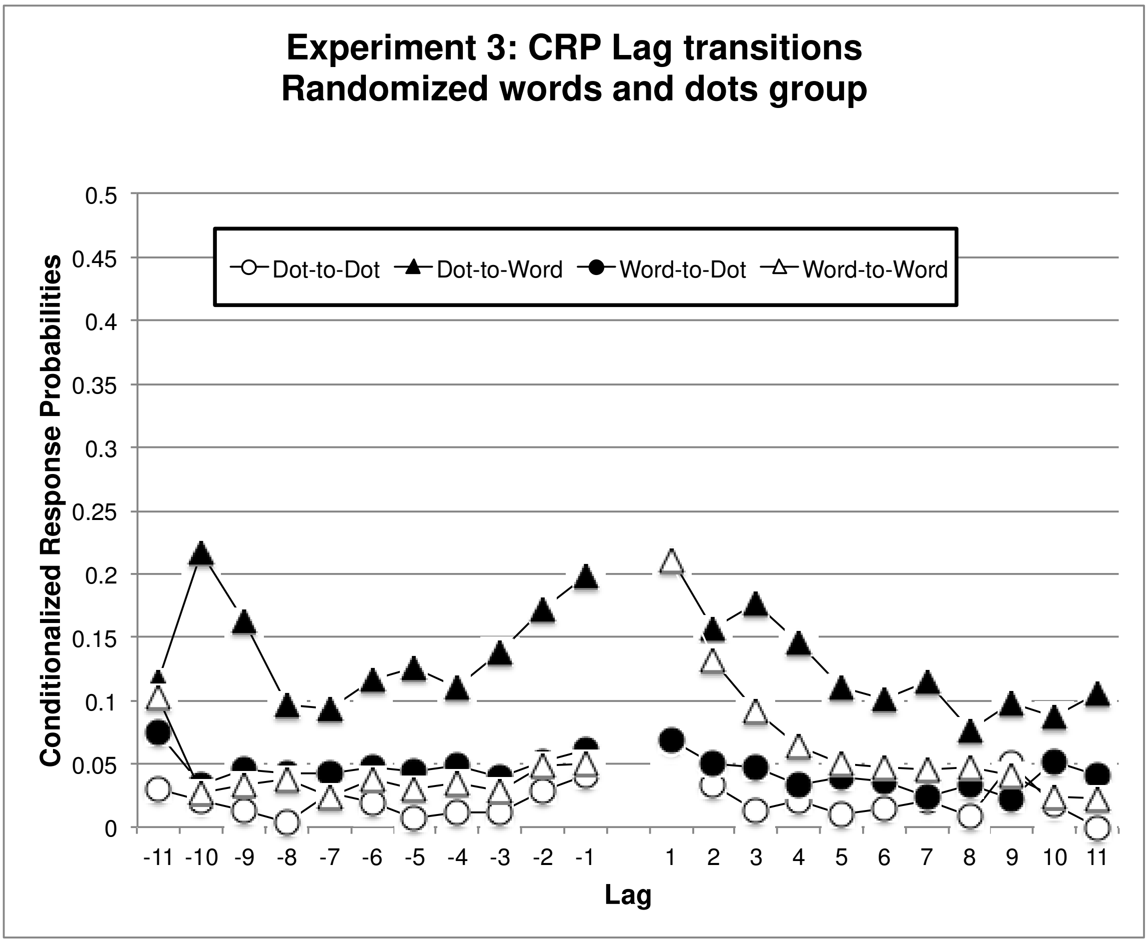

Experiment 3: CRP Lag transitions
Randomized words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilities
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10
11
Lag

Figure 10

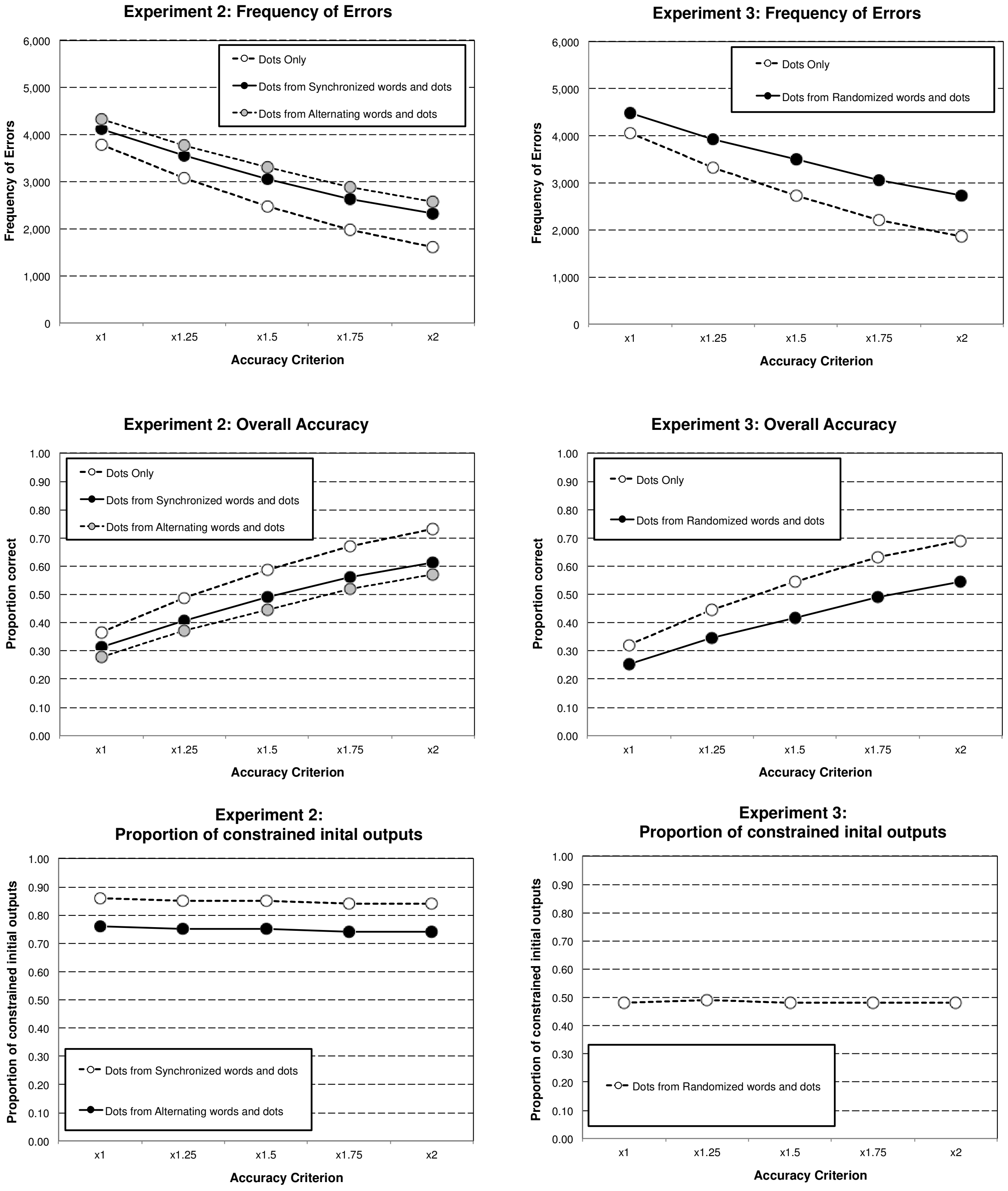
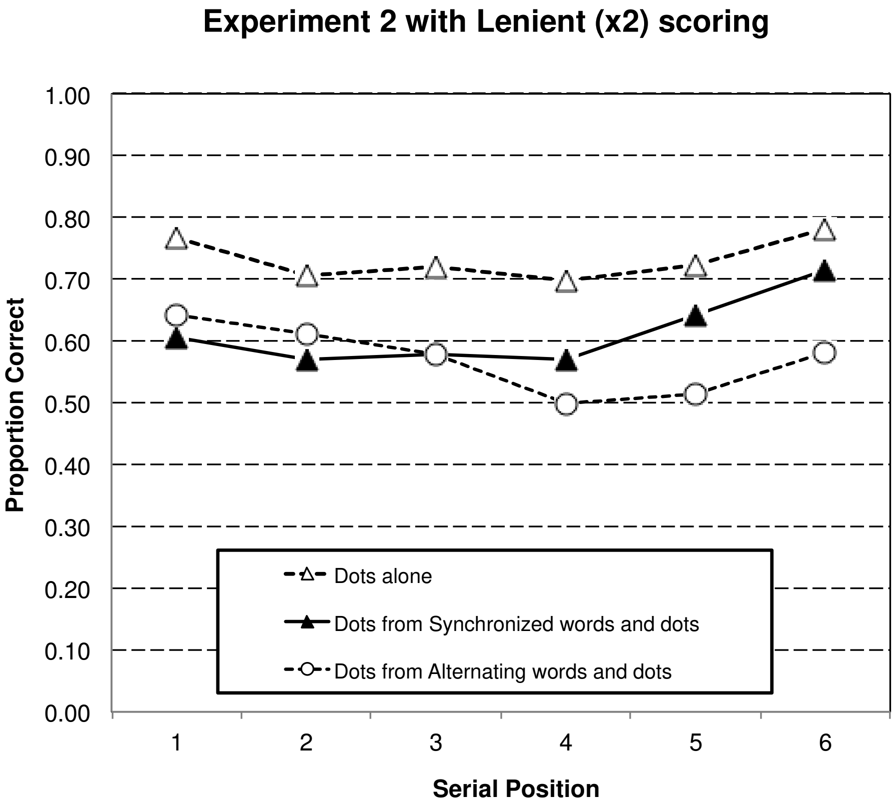

Experiment 2: Frequency of Errors
Frequency of Errors
6,000
5,000
4,000
3,000
2,000
1,000
0
Dots Only
Dots from Synchronized words and dots
Dots from Alternating words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion
Experiment 3: Frequency of Errors
Frequency of Errors
6,000
5,000
4,000
3,000
2,000
1,000
0
Dots Only
Dots from Randomized words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion
Experiment 2: Overall Accuracy
Proportion correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Dots Only
Dots from Synchronized words and dots
Dots from Alternating words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion
Experiment 3: Overall Accuracy
Proportion correct
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Dots Only
Dots from Randomized words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion
Experiment 2:
Proportion of constrained inital outputs
Proportion of constrained initial outputs
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Dots from Synchronized words and dots
Dots from Alternating words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion
Experiment 3:
Proportion of constrained inital outputs
Proportion of constrained initial outputs
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Dots from Randomized words and dots
x1
x1.25
x1.5
x1.75
x2
Accuracy Criterion

Figure 11


Experiment 2 with Lenient (x2) scoring
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion Correct
1
2
3
4
5
6
Serial Position
Dots alone
Dots from Synchronized words and dots
Dots from Alternating words and dots

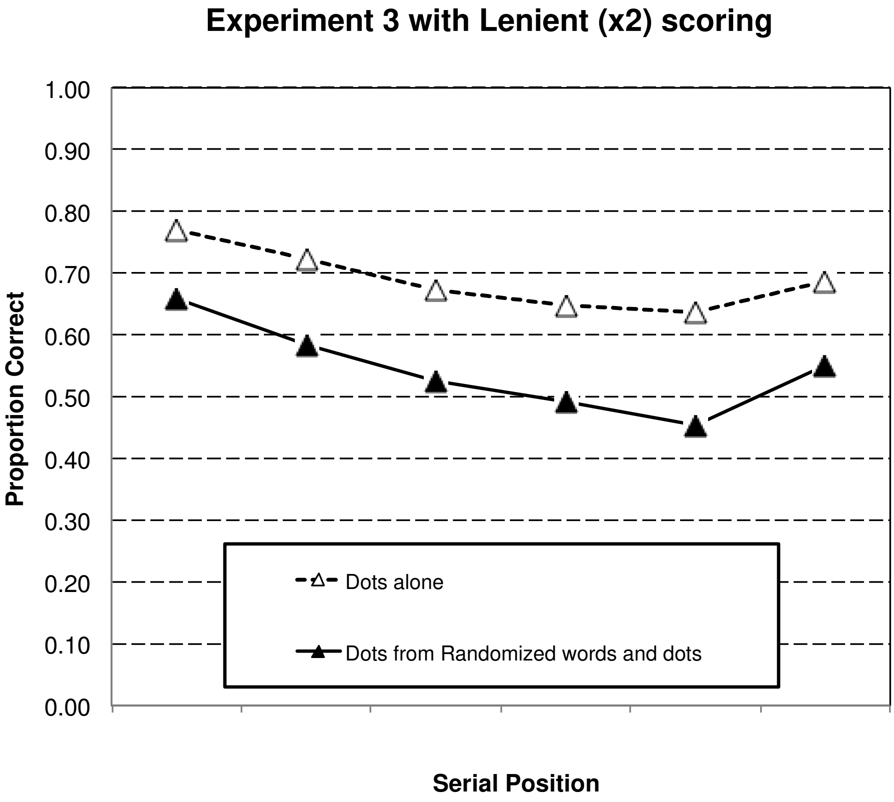
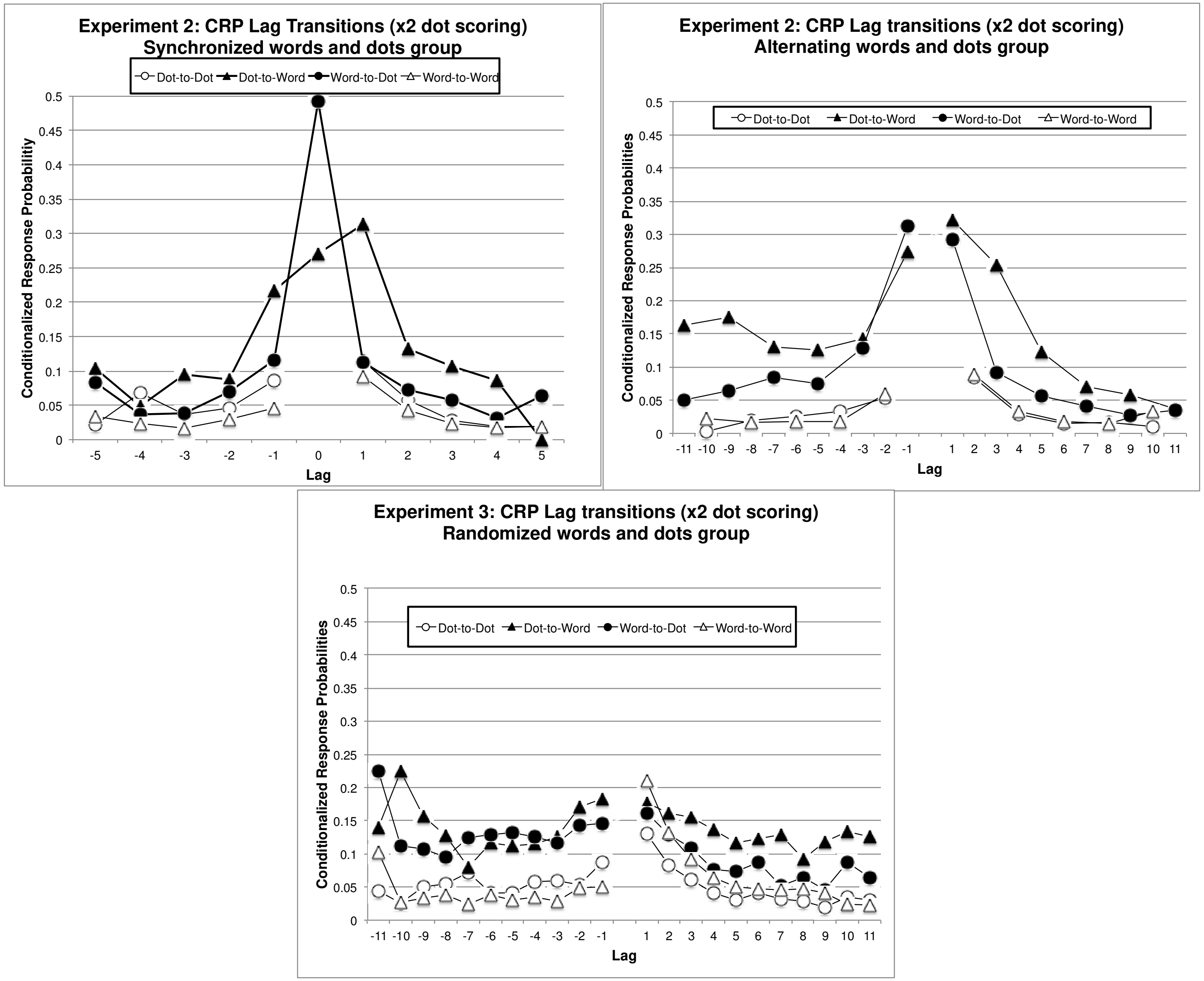

Experiment 3 with Lenient (x2) scoring
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion Correct
Serial Position
Dots alone
Dots from Randomized words and dots

Figure 12


Experiment 2: CRP Lag Transitions (x2 dot scoring)
Synchronized words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilitiy
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
-5
-4
-3
-2
-1
0
1
2
3
4
5
Lag
Experiment 2: CRP Lag transitions (x2 dot scoring)
Alternating words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilities
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10
11
Lag
Experiment 3: CRP Lag transitions (x2 dot scoring)
Randomized words and dots group
Dot-to-Dot
Dot-to-Word
Word-to-Dot
Word-to-Word
Conditionalized Response Probabilities
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
-11
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
1
2
3
4
5
6
7
8
9
10
11
Lag

Appendix A

Analyses of the IFR serial position curves from Experiment 1, shown in Figure 3, using all the data with FR scoring. The performance of each of the single-modality groups (auditory-verbal and visuo-spatial) was compared to the verbal and visuo-spatial performance of the parallel presentation group respectively. At each list length, the IFR data were subjected to a 2(group: either auditory-verbal and parallel presentation or visuo-spatial and parallel presentation) x *n* (serial position 1…*n*) mixed ANOVA where *n* is the list length.

List Length	Main Effects		Interaction
	Group	Serial Position	
Verbal			
2	*F*(1,38) = .159, *MSE* = .003, η^2_p = .004, *p* = .692	*F*(1,38) = .734, *MSE* = .003, η^2_p = .019, *p* = .397	*F*(1,38) = 1.65, *MSE* = .003, η^2_p = .042, *p* = .206
4	*F*(1,38) = .007, *MSE* = .036, η^2_p < .001, *p* = .934	***F*(3,114) = 6.60, *MSE* = .010, η^2_p = .148, *p* = .001**	*F*(3,114) = .558, *MSE* = .010, η^2_p = .014, *p* = .664
8	*F*(1,38) = .775, *MSE* = .052, η^2_p = .020, *p* = .384	***F*(7,266) = 24.4, *MSE* = .040, η^2_p = .391, *p* < .001**	***F*(7,266) = 2.05, *MSE* = .040, η^2_p = .051, *p* = .049**
16	*F*(1,38) = 3.65, *MSE* = .024, η^2_p = .088, *p* = .064	***F*(15,266) = 85.5, *MSE* = .021, η^2_p = .692, *p* < .001**	***F*(15,266) = 2.09, *MSE* = .021, η^2_p = .052, *p* = .009**
Visuo-Spatial			
2	*F*(1,38) = .127, *MSE* = .048, η^2_p = .003, *p* = .723	*F*(1,38) = .639, *MSE* = .024, η^2_p = .017, *p* = .429	*F*(1,38) = .005, *MSE* = .024, η^2_p < .001, *p* = .942
4	*F*(1,38) = 1.55, *MSE* = .071, η^2_p = .039, *p* = .221	*F*(3,114) = 1.97, *MSE* = .031, η^2_p = .049, *p* = .123	*F*(3,114) = .153, *MSE* = .031, η^2_p = .004, *p* = .928
8	*F*(1,38) = 3.74, *MSE* = .056, η^2_p = .090, *p* = .061	***F*(7,266) = 7.94, *MSE* = .025, η^2_p = .173, *p* < .001**	*F*(7,266) = .960, *MSE* = .025, η^2_p = .025, *p* = .461
16	***F*(1,38) = 14.1, *MSE* = .060, η^2_p = .271, *p* = .001**	***F*(15,266) = 14.7, *MSE* = .019, η^2_p = .279, *p* < .001**	*F*(15,266) = 1.32, *MSE* = .019, η^2_p = .034, *p* = .182

Note: significant main effects and interactions are presented in bold