Perceptual-Gestural Mismapping in Serial Short-Term Memory: The Impact of Talker Variability

Robert W. Hughes, John E. Marsh, & Dylan M. Jones

School of Psychology, Cardiff University, Cardiff, UK.

RUNNING HEAD: Perceptual-Gestural Mismapping in Serial Recall

Correspondence: Robert W. Hughes, School of Psychology,

Cardiff University, Cardiff CF10 3AT, United Kingdom,

Phone: (+44) 029 20 876904, Fax: +44 029 20874858,

email: HughesRW@cardiff.ac.uk

Perceptual-Gestural Mismapping in Serial Short-Term Memory: The Impact of Talker Variability

The disruptive impact of talker-variable (e.g., alternating female-male) lists on serial recall was examined. We tested the novel hypothesis that this *talker variability effect* arises from the tendency for perceptual streaming (by voice) to partition the list into two sub-sequences such that the perception of order is in conflict with the formation of a sequence-output plan that remains faithful to the canonical order of the items. The hypothesis was supported by three convergent lines of evidence: Factors known to promote partitioning of items by voice accentuate the effect (Experiments 1 and 2); talker variability combines non-additively with phonological similarity, consistent with the view that both variables disrupt sequence output-planning (Experiment 3); and whereas tasks that require ordered recall—and hence the assembly of a sequence-output plan—show the effect, tasks requiring only item memory do not (Experiments 4 and 5). The results are consistent with the view that serial short-term memory reflects the parasitic use of sequencing processes embodied within general-purpose perceptual input-processing and gestural output-planning systems and are problematic for an item-decay based approach or an item-distinctiveness/attentional-resource account.

KEYWORDS: Short-Term Memory; Talker Variability; Serial Recall; Perceptual-Gestural Account; Embodied Cognition.

The capacity to retain and reproduce a sequence of events over the short-term has long commanded a great deal of interest on the grounds that coherent sequential behaviour is involved in most, if not all, goal-driven activities (e.g., Lashley, 1951). Classically, accounts of serial short-term memory phenomena have been centred at the item level and assume that an understanding of serial behavior will flow from knowledge of item-level properties such as the rate of item decay or/and the structural (e.g., phonological) similarity of one item to another (e.g., Baddeley, 1986, 2007; Farrell & Lewandowsky, 2002; Nairne, 1990; Neath, 2000). A more recently-emerging view—the perceptual-gestural view—focuses on factors that operate at a level superordinate to the item, at the level of sequence formation, both at input (particularly in the auditory modality, in the formation of streams), and at motor output planning in the formation of a sequence of subvocal gestures (Hughes & Jones, 2005; Jones, Macken, & Nicholls, 2004; Jones, Hughes, & Macken, 2006, 2007; Woodward, Macken, & Jones, 2008; see also Buchsbaum & D'Esposito, 2008; Wilson & Fox, 2007). The overarching goal of the present research was to further examine the perceptual-gestural view using the item-based approaches as theoretical counterpoints.

In the present study, interest centres upon a setting which, we hypothesize, may be characterized as one in which there is a poor mapping between auditory perceptual organization and the sequential motor plan (and its eventual output), namely, the talker variability effect in serial recall. This effect refers to the impairment produced in auditory serial recall when successive to-be-remembered items are presented in different voices (Greene, 1991). We test the hypothesis that the effect is the result of the formation of voice-based sub-streams (*auditory streaming*, cf. Bregman, 1990) such that the perceived order of the items (within the sub-streams) is incompatible with the requirement to reproduce the list in serial-temporal order.

Theoretical Approaches to Serial Short-Term Memory

Current understanding of serial short-term memory is based predominantly on the serial recall paradigm in which, typically, a familiar set of verbal items (e.g., the digits *1-8*) is presented in an unfamiliar sequence and participants are asked to reproduce the list in strict serial order (Conrad, 1964, Baddeley, 1966). Classically, explanations of serial recall performance have tended to focus on the properties of the individual items comprising the list. For instance, according to what Nairne (2002) termed the standard (decay-rehearsal) model of verbal short-term memory—best exemplified perhaps by the phonological loop model (Baddeley, 1986, 2007; Baddeley & Hitch, 1974; but see also Atkinson & Shiffrin, 1968; Cowan, 1999)—verbal items are assumed to enter a passive, bespoke, store dedicated to the temporary retention of phonologically-coded traces of each item (cf. the *phonological store;* Baddeley, 1986, 2007). Items in the store decay within about 2 s unless refreshed by covert verbal rehearsal (e.g., Repovs & Baddeley, 2006; Schweickert & Boruff, 1986) but are also susceptible to mutual interference by virtue of their structural (e.g., phonological) similarity to other items (e.g., Baddeley, 1986).

Interference-by-item-similarity also serves as the core explanatory construct in another broad class of theory, namely, that based on item-distinctiveness (e.g., Brown, Neath, & Chater, 2007; Nairne, 1990; Neath, 2000; Farrell & Lewandowsky, 2002). For example, according to the feature model (Nairne, 1990; Neath, 2000) serial recall performance is assumed to bear a simple positive relationship to the distinctiveness—and hence immunity from being overwritten—of the items in a serial recall list in terms of both their modality-dependent features (e.g., pitch) and their modality-independent features (those that do not vary with modality of presentation, e.g., phonology, semantics).

We have suggested recently, however, that appealing to mechanisms that impact upon the assembly and maintenance of sequences, not each item, may prove a more fruitful

approach to understanding performance in serial short-term memory tasks (e.g., Jones et al., 2004; Woodward et al., 2008). An important feature of the typical serial recall study is that the burden of processing falls upon reproducing the order of the items and not upon knowing their individual identities (e.g., Baddeley, 1966): A familiar closed set of items is typically used on each trial (e.g., permutations of *1-8*) and hence the identity of the individual items is known before the list is presented; effectively, therefore, the key task is to retain and reproduce the unfamiliar order in which that familiar item-set has been presented (hereafter we refer to this typical closed-set procedure as *pure* serial recall). Although some serial recall studies employ an open pool of items (see, e.g., Poirier & Saint-Aubin, 1996) where there is also a burden on remembering what items were presented, critically, the four historically and theoretically most important serial recall phenomena are ones that are found in the pure variant of the task (see, e.g., Baddeley, 1990): the *phonological similarity effect* (e.g., Baddeley, 1966), the *articulatory suppression effect* (e.g., Murray, 1968), the *word-length effect* (e.g., Baddeley, Thomson, & Buchanan, 1975) and the *irrelevant sound effect* (Colle & Welsh, 1976).

An alternative means by which serial recall performance has begun to be construed, therefore, is in terms of the parasitic use of general-purpose perceptual and motor-planning processes that operate at the level of the sequence, not each item (e.g., Hughes & Jones, 2005; Jones et al., 2004, 2006, 2007; Woodward et al., 2008; for a similar view based on neuroscientific evidence, see Buchsbaum & D'Esposito, 2008). On this *perceptual-gestural* view, serial recall reflects, in large part, a dynamic, active, process of converting the incoming sequence into gestural form (articulatory in the case of verbal items). In contrast to the standard model, this assembly of verbal items into an articulatory form is not in the service of offsetting the forgetting of individual decay-prone items residing passively in a bespoke store (e.g., Baddeley, 1986). Rather, the process of speech- (or more generally,

motor-) programming is co-opted because its inherent sequentiality—supplemented by a range of paralinguistic speech habits such as co-articulation, intonation, and prosody—acts as a surrogate for those aspects of language such as syntax, grammar, and semantics that usually constrain item order in a normal sentence but which have, by design, been stripped from the serial recall list (Macken & Jones, 2003; for an early precursor of this 'parasitic' view of serial recall performance, see Reisberg et al., 1984). That is, the speech-planning machinery is exploited opportunistically to graft sequential constraints into an artificially impoverished verbal sequence. Accordingly, within this framework, explanations are sought by recourse to sequence-level factors: Performance reflects largely the efficacy with which a fluent *sequence* of gestures can be assembled and rehearsed (sub-vocally) rather than the integrity of stored item representations.

Whilst the gestural component of the perceptual-gestural view applies equally to visually and auditorily presented lists, the perceptual component applies mainly to the auditory domain and draws upon Bregman's (1990) revolutionary ideas regarding *auditory scene analysis*: the partitioning of the mixture of pressure variations reaching the ears into discrete mental descriptions (streams) of each independent sound source contributing to that mixture (Bregman, 1990, 1993; see, e.g., Hughes & Jones, 2005; Jones et al., 2004; Nicholls & Jones, 2002). Of particular interest in the context of serial recall and especially in relation to the present research is *sequential streaming* whereby the perceptual system must determine whether or not temporally successive auditory stimuli are emanating from the same environmental source, a task accomplished by exploiting a host of factors embodied by Gestalt grouping principles such as spectral similarity and good continuation (for an overview, see Bregman, 1993; see also Warren, 1999). One key consequence of this process is that the perception of order has been found to be relatively good for a succession of acoustically-changing stimuli that share some more fundamental common ground (or

carrier) and hence are still assigned to the same stream (e.g., different words spoken in the same voice). Conversely, order perception is poor for successive items that lack a common ground and which will tend not, therefore, to be assigned to the same stream (e.g., different words spoken in different voices; e.g., Bregman & Campbell, 1971; Warren, Obusek, Farmer, & Warren, 1969).

The perceptual-gestural view has already accrued some support in the context of a number of serial recall phenomena classically attributed to item-level constructs, including the *irrelevant sound effect* (Hughes & Jones, 2005; Jones et al., 2004), the *phonological similarity effect* (Jones et al., 2004, 2006), the *suffix effect* (Nicholls & Jones, 2002), and *linguistic familiarity* effects (Woodward et al., 2008). Of interest in the present article is the *talker variability effect* in serial recall: the impairment of auditory serial recall when successive items are presented in different voices (Greene, 1991; see also Goldinger, Pisoni, & Logan, 1991; Martin, Mullennix, Pisoni, & Summers, 1989; Nygaard, Sommers, & Pisoni, 1995).[1] For example, Greene (1991) found that the serial recall of permutations of the digits *1-8* was depressed when the items were presented in alternating female-male voices compared to the conventional, single-voice, mode of presentation. It is worth noting that Greene (1991) employed talker variability as a tool for studying the suffix effect (e.g., Crowder & Morton, 1969) and the talker variability effect itself was a subsidiary concern. The current study is the first, therefore, to utilize the talker variability effect in its own right as a device with which to examine competing approaches to serial short-term memory. In particular, we seek to show how the phenomenon serves to reveal the roles of both auditory perceptual organization and gestural-sequencing processes in serial recall: We hypothesize that the phenomenon is best understood in terms of a disharmony between obligatory-perceptual and deliberate-gestural sequential organization processes.

Competing Accounts of the Talker Variability Effect

Perceptual-Gestural Mismapping Account

From the standpoint of our perceptual-gestural framework, we suggest that the talker variability effect reflects a mismapping between two incompatible sequential organizations, one arising from an obligatory auditory perceptual organization process and the other from a deliberate gestural sequence-output planning strategy. As noted, a single-voice list conveying a succession of different items is an excellent example of coherent change-on-a-common-ground and thus the products of order perception are isomorphic with the items' objective temporal order. In a talker-variable list, however, we suppose that the list's perceptual coherence is greatly diminished, resulting in a perceptual-gestural mismapping. Indeed, when the same two voices (e.g., male and female) alternate in a list (Greene, 1991), it is likely that the items spoken by each voice (i.e., non-adjacent items) will perceptually cohere more readily than temporally successive items due to grouping by similarity of frequency and timbre (cf. Bregman & Campbell, 1971; Carlyon, Cusack, Foxton, & Robertson, 2001). According to the perceptual-gestural view, therefore, the talker variability effect reflects a difficulty in the process of assembling into an articulatory sequence output-plan incoming items whose order—based on the products of obligatory perceptual sequencing processes—maps relatively poorly onto the requirement to assemble the items in their true temporal order.

Item-decay (Standard Model) Accounts

From the perspective of the standard model, talker variability effects in serial recall have been explained in essentially the same manner as the word-length effect (e.g., Baddeley et al., 1975): The increased time taken to encode or/and rehearse talker-variable items—just as with long compared to short words—impairs recall by delaying the opportunity to refresh decay-prone items residing in a bespoke verbal (i.e., phonological) store. For example,

Martin et al. (1989) suggest that talker variability may impose a delay in encoding items into a dedicated verbal short-term store due to a speech normalization process: Talker-variable lists impose a greater burden on the process of discarding indexical information such as the pitch and timbre of the particular speaker's voice in order to yield abstract, canonical, linguistic (i.e., phonological) item representations (see, e.g., Joos, 1948; Magnusson & Nusbaum, 2007). These "increased capacity demands needed for encoding reduce the available resources needed for subsequent rehearsal of the items" (Martin et al., 1989, p. 677). A second item-decay account supposes that the encoding delay is due to an obligatory process of incorporating the indexical voice information rather than discarding it, a process which would again be under greater duress the greater the number of voices in a list (Goldinger et al., 1991; Nygaard et al., 1995). Moreover, in this latter account, the additional information about voice incorporated into each representation would increase "the total amount of information to rehearse per unit time" (Goldinger et al., 1991, p. 159) thereby further exacerbating item decay.

It should be noted that the studies of Goldinger et al. (1991), Martin et al. (1989), and Nygaard et al. (1995) involved a relatively long list-length (10 items), an open pool of words, and a free-output procedure (whereby the outputted items must ultimately correspond to their original serial positions but may be output in any order; see, e.g., Tan & Ward, 2007). In this setting, the effect is only found for early list items whereas it is apparent throughout most of the serial position curve in pure serial recall (Greene, 1991). Thus, whilst we do not assume that these authors (e.g., Martin et al., 1989) would have necessarily applied their item-decay accounts to the effect later found in pure serial recall (Greene, 1991), for current purposes the important point is that, logically, generalizing the accounts to the pure serial recall setting is entirely consistent with the standard model. This follows because the word-length effect (e.g., Baddeley et al., 1975)—the key phenomenon

motivating the concept of item-decay within the standard model, and, in particular, the item-decay accounts of the talker variability effect—is an effect found in pure serial recall (e.g., with closed sets of digits; Ellis & Hennelley, 1980; Murray & Jones, 2002; or a closed set of familiar words; Baddeley et al., 1975).

Item-Distinctiveness/Attentional-Resource Account

The basic talker variability effect seems problematic for item-distinctiveness accounts of short-term memory (e.g., Brown, Neath, & Chater, 2007; Farrell & Lewandowsky, 2002; Nairne, 1990; Neath, 2000). Such models predict straightforwardly that the greater inter-item distinctiveness provided by a talker-variable list at the level of modality-dependent features (pitch, timbre) should lead to better, or at least not poorer, performance: Each item in a talker-variable list should be less prone to interference from (e.g., through overwriting by) its successor because there is less structural overlap (or greater novelty; Farrell & Lewandowsky, 2002) than is the case in a single-voice list. However, a possibility open to some variants of the item-distinctiveness approach such as the feature model (e.g., Neath, 2000) is to appeal to the same device by which that model explains the changing-state irrelevant sound effect in serial recall (Jones, Madden, & Miles, 1992). That to-be-ignored irrelevant changing-state irrelevant sound (“*F R X…*”) is markedly more disruptive of serial recall than steady-state sound (“*F F F…*”) is simulated in the feature model by decreasing the value of an attention parameter (‘a’) that acts to depress the model’s overall performance (Neath, 2000). Thus, it might be supposed that talker variability—like changing-state irrelevant sound—draws upon some general attentional resource (cf. Kahneman, 1973) thereby impairing performance of any attention-demanding task such as serial recall (hereafter the ‘attentional-resource account’). The present series of experiments sought to examine the perceptual-gestural based account of

the talker variability effect, contrasting its predictions with the item-decay accounts and an attentional-resource account.

Experiment 1

Experiments 1-3 follow Greene's (1991) methodology and examine the talker variability effect in pure serial recall using a closed set of either digits (Experiments 1 and 2) or letters (Experiment 3) which were to be recalled in strict serial order. In Experiment 1, we test a prediction that is unique to the perceptual-gestural mismapping account by capitalizing on a particular characteristic of auditory sequential stream segregation, namely, "auditory stream biasing": If a sequence of alternating high (H) and low (L) tones (HLHLHL) is preceded by a succession of either H or L tones (e.g., LLLLLHLHLHL), the partitioning of the alternating sequence into two separate low-tone and high-tone streams occurs more readily (Anstis and Saida, 1985; Beauvois & Meddis, 1997). This is because the lead-in L tones serve to establish a stable stream into which the L tones in the following alternating sequence can be incorporated whilst the other, H, tones are "thrown out" to form a distinct stream. Thus, a lead-in facilitates (or "biases") the partitioning of the alternating stimuli by "perceptually capturing" only the same-frequency tones present in the ensuing alternating sequence. We have demonstrated elsewhere that the same principles hold also for speech stimuli (Nicholls & Jones, 2002). In the present experiment, therefore, we sought to promote the perceptual partitioning of an alternating-voice list in a serial recall task by presenting a lead-in which took the form of a countdown ("8, 7, 6…1") spoken in the same rhythm as the ensuing to-be-remembered items and spoken in just one of the two voices making up the ensuing alternating voice list. Our rationale was that if one critical aspect of the talker variability effect is the perceptual incoherence of temporally successive items in an alternating-voice list, any factor that promotes that incoherence

should cause a further impairment of serial recall (i.e., over and above that found with alternating voices without a lead-in).

Another condition involved preceding an alternating-voice list with an alternating-voice lead-in. Again, this type of lead-in should bias the partitioning of alternating-voice items in a to-be-remembered list on the grounds that partitioning takes time to "build up" (e.g., Bregman, 1978; Carlyon et al., 2001). Thus, with pre-exposure to the alternating pattern of voices, partitioning by voice will have begun to be established before the to-be-remembered list begins. Again, therefore, an alternating lead-in should accentuate the disruptive impact of an alternating-voice list. Table 1 provides a list of all six conditions contrasted in Experiment 1. Conditions 1 and 2 represent those required to show the standard talker variability effect (i.e., those without a lead-in). The remaining four conditions represent a factorial combination of lead-in type (single or alternating voice) and list-type (again, single or alternating voice). To summarize, in relation to conditions 1-6 shown in Table 1, the pattern of performance (going from best to worst) predicted by the perceptual-gestural mismapping account is as follows: 1=3=5>2>4=6.

In contrast to the prediction of the perceptual-gestural mismapping account, on the item-decay accounts there is no reason to expect a lead-in to accentuate the talker variability effect. Indeed, logically, these accounts would predict, if anything, a facilitative effect of a lead-in: Pre-exposure to (and hence pre-knowledge of) one of the voices (single-voice lead-in) and particularly to both voices (alternating lead-in) might be expected to ease the burden on the process of voice normalization (Martin et al., 1989) or voice incorporation (Goldinger et al., 1991) when the time comes to encode the identity of the to-be-remembered items thereby resulting in a reduction of the talker variability effect. In this case, item-decay accounts predict the pattern: 3>5>1>6>or=4>2. A more conservative

prediction that seems open to decay-accounts is that lead-ins will simply have no impact on performance.

Given that the psychological mechanism by which the attentional resource represented by parameter 'a' in the feature model (Neath, 2000) is depleted has yet to specified in detail (see, e.g., Jones & Tremblay, 2000), it is not always obvious what predictions might be derived from the attentional-resource account (although see Experiments 2-4). However, one reasonable expectation in relation to Experiment 1 might be that performance should simply be a negative function of the degree of talker variability contained within the lead-in or/and the to-be-remembered list (with the possible additional assumption that talker variability within the list itself would be particularly damaging). If so, the pattern of performance predicted by the attentional-resource account is: 1=or>3>5>2>6>4.

Method

Participants

Twenty-two undergraduates from Cardiff University took part in return for course credits. Each participant reported normal hearing and normal or corrected-to-normal vision.

Apparatus and Materials

The to-be-remembered lists comprised 8 items taken without replacement from the digit-set 1-8. Each item was recorded digitally twice, once in a female voice and once in a male voice (the items within each voice were spoken at an approximately even pitch), and sampled with a 16-bit resolution at a sampling rate of 44.1KHz using *Sound Forge 5* software (Sonic Inc., Madison, WI; 2000). The male and female voices clearly differed from one another on account of their distinct fundamental frequency and timbre. Each item's duration was edited to 250ms. For each to-be-remembered list, the digits were presented in a pseudo-random order with care taken to ensure that there were no more than

two occasions across a given to-be-remembered list on which there was an ascending or descending run of two or more digits (e.g., 2-3 or 7-6) and that there were no runs of 3 or more digits. This was also the case for non-adjacent items (e.g., those in positions 1 and 3) so that in alternating-voice lists there were no more than two 2-digit runs within a given voice in a given list. The to-be-remembered list (and lead-in when present) was presented at approximately 65-70 dB(A) over stereo headphones with an inter-stimulus interval (ISI; offset to onset) of 100ms giving an item presentation rate of 1 item/350ms. Although this is a faster presentation rate than used in Greene's (1991) study (1 item/s), it was adopted here to promote the chances of bringing into relief the contribution of perceptual organization processes to the effect: It is well established that the perceptual incoherence—and hence partitioning—of temporally successive sounds alternating in frequency (such as the male- and female-spoken items used here) is a positive function of the rate at which they are presented (e.g., Bregman, 1990; van Noorden, 1975; Warren, Obusek, Farmer, & Warren, 1969). Note however that the rate we adopted is still not far removed from the 1 item/500ms rate often used in serial recall studies (e.g., Baddeley et al., 1984; Farrell & Lewandowsky, 2003; Lewandowsky & Farrell, 2008; see also *Supplementary Experiment* in *Results* section of Experiment 1). The stimuli were presented using the *SuperLab* software (Cedrus Corporation).

Materials for 6 conditions were assembled (see Table 1): In all conditions in which the to-be-remembered list was presented in a single voice (i.e., Single, Single-Single, and Alt-Single conditions), half the lists were spoken entirely in the male voice and half were spoken entirely in the female voice. In the other three conditions—Alt, Alt-Alt, and Single-Alt—the list was presented in an alternating female-male fashion with half the lists starting with a female-spoken item and half starting with a make-spoken item. In conditions involving a single-voice lead-in, a countdown was presented either in the same voice as the

ensuing single-voice list (Single-Single) or, for the Single-Alt condition, in the same voice as that conveying the second, fourth, sixth, and eighth items of the ensuing alternating-voice list. In the Alt-Single condition, the countdown was presented in alternating female-male voices starting with the female voice if the to-be-remembered list was female-spoken and with the male voice if the to-be-remembered list was male-spoken. In the Alt-Alt condition, the pattern of alternation always continued unbroken into the ensuing to-be-remembered list.

Design

The experiment had a repeated-measures design with three factors: Lead-in (with three levels: no lead-in, single-voice lead-in, and alternating-voice lead-in), List-type (with two levels: single-voice and alternating-voice), and Serial position (eight levels). Each participant undertook 84 experimental trials divided into two blocks. One block—the 'with lead-ins block'—comprised 56 experimental trials made up of: 14 Alt-Alt trials (7 in which the to-be-remembered list started with a female item and 7 in which it started with a male item); 14 Alt-Single trials (7 in which the to-be-remembered list was female-spoken and 7 in which it was male-spoken); 14 Single-Single trials (7 in which the to-be-remembered list was female-spoken and 7 in which it was male-spoken); and 14 Single-Alt trials (7 in which the to-be-remembered list started with a female item and 7 in which it started with a male item). The four trial-types were presented pseudo-randomly across the 56 trial-block with the constraint that no condition was presented more than twice in succession. The block was preceded by 4 practice trials, one from each of the four conditions. The other block—the 'without lead-ins block'—comprised 28 experimental trials made up of 14 single-voice to-be-remembered lists (7 female, 7 male), and 14 alternating-voice to-be-remembered lists (7 female-first, 7 male-first) preceded by 2 practice trials, one from each condition. The two trial-types were presented pseudo-randomly across the 28 trials with the

constraint that no condition was presented more than twice in succession. The order in which the two blocks were undertaken was counterbalanced across participants.

Procedure

Participants were tested in groups of up to four in a sound-attenuated room with each participant placed in a separate cubicle with its own PC and headphone. Participants were instructed to attempt to recall the to-be-remembered digits in their correct order and to ignore the particular voice(s) conveying the digits. Participants were also told that for one block of trials the spoken list would be preceded by a spoken countdown. They were informed that 100ms following the offset of the last to-be-remembered item of each list, the cue 'RECALL' would appear on the screen and that at this point they should try to write down the items in the correct order on response sheets marked with 8 blank spaces for each trial. Participants were told that they had 15 s to write down the list and that they should do so in a strict left to right fashion such that they should start by writing the item they recalled as having occurred first in position 1, then go on to position 2, and so on. They were instructed to guess if they were uncertain of any of the digits' positions. A 500ms tone was presented over the headphones 13 s into the 15 s recall-period to signal to the participant that the presentation of the first item of the next trial was imminent (in trials with a lead-in, the first item would of course be the first item of the countdown). Including an optional 5 min break between the two blocks, the experiment lasted approximately 45 min.

Results

For Experiments 1-3, the raw serial recall data were scored according to the strict serial recall criterion: To be recorded as correct an item had to be recalled in its original presentation position. Figure 1 shows the percentage of items correctly recalled across the eight serial positions in the six conditions. The pattern of results is clear-cut and can be

unpacked initially into two distinct sets of curves: Replicating the basic talker variability effect, performance in conditions involving an alternating to-be-remembered list (i.e., Alt, Single-Alt, and Alt-Alt; represented by the triangle symbols) was uniformly poorer than for conditions involving a single-voice list (i.e., Single, Alt-Single, and Single-Single; represented by the square symbols). More importantly, the talker variability effect was markedly accentuated by the presence of a lead-in: Performance with alternating-voice lists was particularly poor when those lists were preceded by either an alternating- or single-voice lead-in (Single-Alt and Alt-Alt). The pattern across conditions thus conforms to that predicted by the perceptual-gestural mismapping account—1=3=5>2>4=6—and is at variance with that predicted by the item-decay (3>5>1>6>or=4>2) and attentional-resource accounts (1=or>3>5>2>6>4).

A 2 (List-type) by 3 (Lead-in) by 8 (Serial Position) repeated-measures ANOVA confirmed that the just-described pattern evident in Figure 1 was statistically reliable: First, the replication of the classical serial position curve depicted across all conditions in Figure 1 was reflected in a main effect of serial position $F(7, 147) = 55.27$, $MSE = .06$, $p < .001$. Of more interest, there was a main effect of List-type, $F(1, 21) = 69.83$, $MSE = .07$, $p < .001$, a main effect of Lead-in, $F(2, 42) = 15.87$, $MSE = .01$, $p < .001$, and, most importantly, a significant List-type by Lead-in interaction, $F(2, 42) = 12.19$, $MSE = .02$, $p < .001$, reflecting the fact that the talker variability effect was larger when an alternating list was preceded by a lead-in (of either type). The only other significant effect was an interaction between List-type and Serial position, $F(7, 147) = 18.37$, $MSE = .01$, $p < .001$, possibly reflecting ceiling effects at primacy and recency serving to obscure differences according to list-type. Follow-up simple effects analyses confirmed that all alternating-voice to-be-remembered list conditions produced poorer performance than any of the conditions with a single-voice to-be-remembered list (all comparisons $p < .005$). More

importantly, they also showed that performance was poorer in both the Single-Alt and Alt-Alt conditions than in the Alt condition (both $p < .001$).

Supplementary Experiment

Given that the presentation rate used in Experiment 1 was relatively fast (1 item/350ms), we ran a supplementary experiment—not reported in full here for the sake of space—to check that the same interaction between talker variability and lead-in is found also with a slower rate typical of some serial recall experiments (1 item/750ms, e.g., Divin, Coyle, & James, 2001; Henson, Hartley, Burgess, Hitch, & Flude, 2003; Hughes, Vachon, & Jones, 2007). Other than the presentation rate—which we increased to 1item/750ms by changing the inter-stimulus interval to 500ms—the experiment was essentially identical to the main Experiment 1 except we only included the Single, Alt, Single-Single, and Alt-Alt conditions (given that the main experiment had already demonstrated that a lead-in per se does not disrupt recall and given that there was no difference in the efficacy with which the two types of lead-in accentuated the impairment seen with an alternating-voice list). The same pattern was found: There was a main effect of Serial position, $F(7, 175) = 89.07$, $MSE = .02$, $p < .001$, a main effect of List-type, $F(1, 25) = 19.90$, $MSE = .03$, $p < .001$, no main effect of Lead-in, $F < 1$, but again a significant interaction between Lead-in and List-type, $F(1, 25) = 9.15$, $MSE = .02$, $p < .01$, whereby the talker variability effect was larger with a lead-in than without. As well as providing a useful replication of the novel aspect of the main experiment—the impact of a lead-in on the talker variability effect—the results of this supplementary experiment indicate that using a relatively fast presentation rate to investigate the functional characteristics of the talker variability effect is unlikely to compromise the generalizability of the results.

Discussion

The results of Experiment 1 confirm a prediction that is unique to the perceptual-gestural mismapping account of the talker variability effect: The lead-in (of either type) is assumed to have promoted the perceptual incoherence of adjacent to-be-remembered items—and at the same time promote the perceptual coherence of non-adjacent items—thereby accentuating the mismapping between the order suggested by streaming and the action requirements of the serial recall task. Accordingly, the talker variability effect was significantly larger when an alternating voice list was preceded by a lead-in. At the same time, the pattern of results is at odds with the predictions derived from the item-decay and attentional-resource accounts. Item-decay accounts cannot readily explain how the presence of a lead-in could accentuate the talker variability effect. Indeed, if anything, one might expect that pre-exposure to the attributes (e.g., frequency, timbre) of one (Single-Alt condition) or both of the voices (Alt-Alt condition) conveying the ensuing to-be-remembered list—and particularly being pre-exposed to the temporal pattern of voice-changes (as would be the case in the Alt-Alt condition)—would facilitate the process of either normalizing (Martin et al., 1989) or incorporating (Goldinger et al., 1991) those attributes. Such facilitation should in turn have allowed greater opportunity to refresh decay-prone item-representations via rehearsal and hence reduce the magnitude of the talker variability effect. The opposite pattern was in fact observed.

In relation to an attentional-resource account (e.g., Neath, 2000), this account cannot explain why the impact of an alternating- compared to a single-voice list is greater when preceded by a lead-in. It cannot appeal to the notion that a lead-in per se depleted a general attentional resource because a lead-in (of either type) had no effect when preceding a single-voice list. It is worth noting that this latter feature of the results also allows us to reject the potential argument that the impact of a lead-in on the talker variability results

from the lead-in increasing the effective length of the to-be-remembered list and making the task more difficult (for a similar argument in relation to the phonological similarity effect, see, e.g., Baddeley & Larsen, 2007; but see Jones et al., 2007).

In Experiment 2, we turn to test a prediction that derives more directly from both the item-decay and attentional-resource accounts by varying the number of different voices conveying the to-be-remembered list, a prediction that again contrasts with that which flows from the perceptual-gestural mismapping account.

Experiment 2

The item-decay accounts of the talker variability effect are based on the notion that the variation in voices in a talker-variable list increases the burden on encoding or/and rehearsing each item which in turn compromises recall through increased decay. It follows, therefore, that a strong prediction of these accounts is that the magnitude of the talker variability effect should increase as a function of the number of different voices in the to-be-remembered list. It seems plausible to assume that a general attentional resource (Neath, 2000) would also be depleted to a greater extent the greater the variation in voices across the list. Thus, this account makes the same prediction in this case as the item-decay accounts.

In Experiment 2, therefore, we contrasted conditions in which the to-be-remembered list could be conveyed in a single voice, in two (alternating) voices, or four voices. The four voices comprised the female samples used in Experiment 1 (and 2) and another three ‘voices’ generated by pitch-shifting those female (F) samples down by 3 semi-tones (hereafter: F-), up by 3 semi-tones (F+), and up by 6 semi-tones (F++). The single-voice lists could be conveyed in either one of the four voices whilst the alternating-voice lists involved an alternation between F and F+ (or F+ and F). In the four-voice condition, each eight-item list was conveyed in the following pattern of voices: F F+ F++

F+ F F- F F+ (or its mirror image: F+ F F- F F+ F++ F+ F). We purposefully pitch-shifted the original female voice rather than using the original sets of male and female voices and recording another two additional talkers so that the degree of acoustic difference between each successive pair of adjacent items across the two and four-voice conditions was roughly equal (this was also the reason for choosing the particular pattern of voice-changes used in the four-voice condition). If we had used recordings from four different talkers (or used a different pattern) it would be difficult to know whether any difference found between alternating-voice and four-voice lists was related to the number of different voices across the list or a difference in the acoustic distinctiveness between successive items across the two conditions. For each list-type, given that we have shown that the talker variability effect is more robust when lead-ins are used, each to-be-remembered list was preceded by a lead-in (again, a countdown) in which the voice or pattern of voices conformed to that characterizing the ensuing to-be-remembered list. In short, the item-decay and attentional accounts would predict the following pattern of performance: single-voice > alternating voices > four-voice.

In terms of the perceptual-gestural mismapping account, it is reasonable to expect performance to be poorer in both the alternating- and four-voice conditions than in the single-voice condition given the far greater perceptual incoherence in the two talker-variable conditions. However, in contrast to the item-decay and attentional accounts, the four-voice condition should not produce poorer performance than the alternating-voice condition. In fact, perceptual grouping by voice—and hence perception-action mismapping—should be stronger in the alternating-voice condition than in the four-voice condition. This follows on the grounds that the likelihood of temporally non-adjacent items perceptually “capturing” one another into the same stream would be a function of both their acoustic similarity and the number of times those similar items are encountered (see

Bregman, 1990). Thus, given that in the alternating-voice lists there are six instances in which non-alternating items are in the same voice whereas there are only two such instances in the four-voice list, the propensity for non-adjacent items to perceptually capture one another to form a coherent stream (and hence the degree of perceptual-gestural mismapping) is greater in the alternating- compared to the four-voice list. The pattern of performance predicted by the perceptual-gestural mismapping account, therefore, is as follows: single-voice > four-voices > alternating-voices.

The present experiment also serves as a test of a further possible interpretation of the talker variability effect which, at first glance, bears a strong resemblance to the perceptual-gestural mismapping account: Greene (1991) suggested (but did not directly test) the possibility that when presented with an alternating-voice list, participants may adopt a deliberate, but counterproductive, strategy of grouping (or rehearsing) the to-be-remembered items by voice (cf. Tulving and Colotla, 1970). Although this deliberate-grouping account and the perceptual-gestural mismapping account share an emphasis on the role of sequential organization rather than item-level factors, they are nevertheless distinct: On the perceptual-gestural account, grouping by voice is a product of non-strategic (that is, obligatory) primitive auditory perceptual organization processes, not a deliberate, voluntary, strategy. Moreover, on the perceptual-gestural mismapping account, we assume that this involuntary by-voice grouping impairs the attempt to assemble the items into a rehearsal cohort, not by voice but by-order-of-presentation. In contrast, on the deliberate-grouping account, the locus of the difficulty is at output: having deliberately assembled the items by voice, the participant must somehow re-organize the items in an attempt to reproduce the items in their original temporal order.

The contrast between the two- and four-voice conditions in Experiment 2 should allow us to adjudicate between the perceptual-gestural mismapping account and the

deliberate-grouping account: Whilst an alternating-voice list would potentially readily lend itself to a strategy of deliberately assembling the items into two by-voice groups, it would seem less plausible to suppose that participants would adopt a strategy of grouping items into four groups of two same-voice items. Thus, whereas the perceptual-gestural mismapping account predicts a marked decrement in the four-voice condition (as well as in the alternating-voice condition), it is less clear how the deliberate-grouping account could explain a decrement in the four-voice condition.

Method

Participants

Forty undergraduates from Cardiff University took part in return for course credits. Each participant reported normal hearing and normal or corrected-to-normal vision.

Apparatus and Materials

The apparatus and materials were identical to those used in Experiment 1 except that three new sets of voice samples were generated by pitch-shifting the original female-spoken items down by 3 semi-tones, up by 3 semi-tones, and up by 6 semi-tones (without altering each item's duration) using the 'pitch-shift' function in the *Soundforge 7* software.

Design

The experiment had a repeated-measures design with two factors: List-type (three levels: Single-voice, Alternating-voice, and Four-voice) and Serial position (eight levels). Each participant undertook one block of 84 experimental trials (with an optional break of up to 5 min after 42 trials) made up of 28 Single-voice trials (7 in each voice: F, F-, F+, and F++), 28 Alternating-voice trials in which the F voice alternated with the F+ voice (14 started with the F voice, 14 with the F+voice) and 28 Four-voice trials (14 forming the pattern F F+ F++ F+ F F- F F+ and 14 forming the pattern F+ F F- F F+ F++ F+ F). Each given to-be-remembered list was preceded by a lead-in that conformed to the same voice-

format as that to-be-remembered list. The different trial-types were presented pseudo-randomly across the block with the constraint that none was presented more than twice in succession. There were 8 practice trials before the 84-trial block (one of each of the variety of trial-types just listed). The procedure was identical to that of Experiment 1.

Results

Figure 2 shows serial recall performance across the eight serial positions in the three list-type conditions. It is evident that performance was markedly impaired in both talker-variable conditions compared to the single-voice condition. More importantly, in line with the perceptual-gestural mismapping account, and against the item-decay and attentional-resource accounts, performance was slightly but significantly worse in the alternating-voice condition than in the four-voice condition. A repeated-measures ANOVA revealed a main effect of Serial Position, $F(7, 273) = 120.72$, $MSE = .03$, $p < .001$, and of List-type, $F(2, 78) = 49.87$, $MSE = .02$, $p < .001$. As in Experiment 1, the List-type by Serial Position interaction also reached significance, $F(14, 546) = 5.57$, $MSE = .006$, $p < .001$, again possibly due to differences between conditions being obscured by primacy and recency effects. Planned repeated contrasts showed that performance in the Four-voice condition was significantly poorer than in the Single-voice condition, $F(1, 39) = 63.27$, $MSE = .05$, $p < .001$, and, importantly, that performance in the Alternating-voice condition was significantly poorer than in the Four-voice condition, $F(1, 39) = 5.9$, $MSE = .02$, $p = .02$.

Discussion

The results of Experiment 2 converge with those of Experiment 1 in providing support for the perceptual-gestural mismapping account and are at variance with those predicted by the item-decay and attentional-resource accounts. The additional burden on normalizing (Martin et al., 1989) or incorporating (Goldinger et al., 1991) the indexical

attributes of four voices compared to two should have delayed further the process of encoding items into a verbal short-term store and—in the case of the voice-incorporation account—added to the amount of item information to be rehearsed. Hence, the talker variability effect should have been accentuated by the addition of two further voices. In fact, although we acknowledge that the difference between the two-voice compared to four-voice condition was numerically small, there was a significant effect in the opposite direction: four-voice lists were significantly *better* recalled than two-voice lists. This result also goes against an attentional-resource account: it is difficult to envisage why encountering four voices would deplete general attentional resources to a lesser degree than two voices.

The result is consistent, however, with the perceptual-gestural mismapping account: Although perception of true temporal order (i.e., of the immediately adjacent items) would be impaired by the lack of a coherent common carrier in both talker-variable conditions, the non-adjacent items in the alternating-voice list condition—due to their greater acoustic similarity—would be expected to cohere more readily than was the case in the four-voice condition (see Bregman & Rudnicky, 1975). The results of Experiment 2 also present difficulties for a deliberate grouping-by-voice explanation of the talker variability effect (Greene, 1991): Assuming that participants would not readily be able to deliberately group the items by voice in the four-voice condition, this account seems to encounter difficulties explaining the marked decrement in this condition compared to the single-voice condition. It might at least have been expected on this account that the simpler two-voice grouping would have caused less impairment than a four-voice grouping when it came to re-organizing the items into canonical order at output, an expectation at odds with the pattern obtained.

The results of the series thus far support the contention that one key component of the talker variability effect is an obligatory auditory perceptual organization of the items that conflicts with the true temporal order of the items. The second key aspect of our account is that the locus of the impairment is, ultimately, the gestural sequence output-planning process: The process of assembling the items into a gestural analogue is fed incompatible information regarding the order of the items by an obligatory auditory perceptual organization process. At this juncture, therefore, we turn to begin examining the gestural-planning component of our account, once again using the standard model and item-distinctiveness based approaches as theoretical counterpoints.

Experiment 3

The analytical device of examining whether two or more variables known to independently affect serial recall combine to produce an additive or non-additive effect has played an instrumental role in the development of theories of short-term memory (e.g., Baddeley et al., 1984; Jones et al., 2004; Longoni, Richardson, & Aiello, 1993). For example, the non-additivity of the irrelevant sound effect (e.g., Colle & Welsh, 1976) and articulatory suppression (e.g., Murray, 1968) is taken by both the perceptual-gestural view and the feature model as indicating that they share a functional locus (albeit a different one in the two accounts; see Jones et al., 2004; Neath, 2000). By the same token, that the word-length effect (Baddeley et al., 1975) is additive to the phonological similarity effect (e.g., Baddeley, 1966; Conrad, 1964) has been taken by proponents of the standard model as evidence that these two effects have distinct functional loci (Longoni et al., 1993). In Experiment 3, we examine for the first time the possible interplay between talker variability and the phonological similarity effect on the grounds that the perceptual-gestural view predicts their non-additivity whereas according to the competing accounts they should be additive.

The phonological similarity effect is a benchmark finding in serial recall and refers to the finding that phonologically similar items (*b, d, v*….) are more difficult to serially recall than phonologically dissimilar items (*f, r, q*…). Within the phonological loop model—the most successful instantiation of the standard model (Nairne, 2002)—this effect was the main catalyst for the notion of a passive phonological store and has been considered thereafter as the chief empirical signature of its action: "because the phonological store relies purely on a phonological code; similar codes present fewer discriminating features between items, leading to impaired retrieval and poorer recall" (Baddeley, 1992, p. 9). In support of the perceptual-gestural view, however, more recent evidence indicates that the phonological similarity affects the articulatory rehearsal process, perhaps through its promotion of speech-planning errors (Ellis, 1980; Jones et al., 2004, 2006), not an impaired capacity to discriminate similar items in a separate, passive, store. An attribution of the phonological similarity effect to the rehearsal process had previously been rejected on the grounds that the effect was still found when rehearsal was precluded by articulatory suppression so long as the items gained obligatory access to the phonological store by being presented auditorily (Baddeley et al., 1984). However, more recent studies have suggested that the phonological similarity effect found with auditory presentation under suppression is better explained in terms of the parasitic use of pre-phonological (i.e., acoustic) auditory perceptual organization processes than retrieval from a post-perceptual phonological store (Jones et al., 2004, 2006, 2007; but see Baddeley & Larsen, 2007).

The different accounts of the phonological similarity effect held by the standard model (e.g., Baddeley, 2007) and the perceptual-gestural framework (e.g., Jones et al., 2004) provides a further means of adjudicating between the item-decay and perceptual-gestural mismapping accounts of the talker variability effect. A key finding that has been

taken as support for the fractionation of the phonological loop into a passive post-perceptual phonological store and an articulatory rehearsal process is that the phonological similarity effect combines additively with the word-length effect: "…articulatory duration and the phonemic confusability of items to be remembered exert additive and independent effects upon performance in immediate serial recall, and hence [this shows] that they reflect distinct components of the working-memory system (Longoni et al., 1993, p. 14). Specifically, within the phonological loop model, whereas the phonological similarity effect reflects a confusion-during-retrieval between similar phonological item-traces in the passive store, the word-length effect reflects a race between articulatory rehearsal and item decay.

The rationale for Experiment 3 was as follows: Given that the item-decay accounts of the talker variability effect appeal to the same decay-rehearsal mechanism as used to explain the word-length effect, they also predict that phonological similarity and talker variability should exert independent (i.e., additive) effects. In contrast, on the perceptual-gestural view, although talker variability and phonological similarity may affect the sequence output-planning process in rather different ways—the former by making it difficult to initially assemble the items into a rehearsal cohort (or speech-plan), the latter by promoting speech-planning errors during the cyclical elaboration and execution of that speech-plan (for further discussion, see Jones et al., 2004)—the important point for present purposes is that they nevertheless both affect the sequence output-planning process. This view therefore predicts that the two effects will interact (i.e., will be non-additive): The phonological similarity effect should be smaller with talker-variable lists because the speech-planning process will have already been corrupted to some extent by talker variability.

Item-distinctiveness accounts of serial short-term memory, as with the standard model, also construe the phonological similarity effect as being due to the greater structural overlap between phonologically similar items making them more difficult to discriminate at retrieval. For example, in the feature model, successive items will overwrite one another's features to the extent that they are phonologically similar (Nairne, 1990; Neath, 2000). Regardless of whether proponents of the feature model would in fact appeal to the attentional parameter to account for the talker variability effect as we have speculated, the item-distinctiveness approach generally cannot ascribe the talker variability as well as the phonological similarity effect to the concept of item distinctiveness. This follows because, as noted earlier, the concept of item distinctiveness would, contrary to the data, predict a facilitative, not a negative, effect of talker variability. Thus, item-distinctiveness accounts make the same prediction as the item-decay accounts: phonological similarity and talker variability effects should combine additively.

Experiment 3, therefore, required the serial recall of 6 letters which could either be phonologically similar or dissimilar and these two types of list were presented either in a single voice or in alternating voices. Whilst the perceptual-gestural mismapping account predicts an interaction between the two variables, both the standard model and the item-distinctiveness based accounts predict that their effects should be additive.

Method

Participants

Twenty undergraduates from Cardiff University took part in return for course credits. Each participant reported normal hearing and normal or corrected-to-normal vision.

Apparatus and Materials

Four list-types were generated representing a factorial combination of phonologically similarity and talker variability. Each list comprised 6 letters that were

either phonologically dissimilar (*k, q, h, y, r, m*) or similar (*p, d, t, v, b, g*). Female and male-spoken versions of the items (the same male and female that produced the digits in Experiment 1) were recorded, edited, and presented in the same manner as in Experiment 1.

Design

A repeated-measures design was used with 3 factors: Serial position; Phonological similarity (similar vs dissimilar); and Voices (single vs alternating). There were 2 blocks, one comprising 20 phonologically dissimilar lists and the other 20 phonologically similar lists. The order in which these 2 blocks were undertaken was counterbalanced across participants. Within each block, 10 lists were presented in a single voice (5 female; 5 male) and 10 in alternating female-male voices (5 starting with a female-spoken item; 5 starting with a male-spoken item). Within each block, these 4 'voice-type' lists (single-female, single-male, alternating, female first; alternating, male first) were intermixed pseudo-randomly with the constraint that no type of list was presented twice in succession.

Procedure

The procedure was the same as Experiment 1 except for the following details. Following the procedure of Henson, Norris, Page, and Baddeley (1996), before each block, the 6 consonants to be used in that block were presented in a circle on the screen for 2 minutes to allow the participant to familiarize themselves with the closed item-set. They were also given 4 practice trials (one corresponding to each voice-type list) before the experimental trials. The experiment took approximately 25 min.

Results

Figure 3 shows serial recall performance across the six serial positions in the four conditions. It is apparent that whereas with single voice lists there was a very large phonological similarity effect, the effect is attenuated markedly with alternating voice lists.

Thus, the data exhibit a non-additivity of phonological similarity and talker variability in line with the perceptual-gestural mismapping account but at variance with the predictions of competing accounts. The results of a repeated-measures ANOVA supported the foregoing observations. There was a main effect of Serial position, $F(5, 95) = 83.39$, $MSE = 0.03$, $p < .001$, a main effect of Phonological similarity, $F(1, 19) = 68.51$, $MSE = 0.11$, $p < .001$, a main effect of Voice, $F(1, 19) = 16.15$, $MSE = 0.04$, $p < .001$, and, most importantly, a significant interaction between Phonological similarity and Voice, $F(1, 19) = 13.16$, $MSE = 0.03$, $p < .01$. The interactions between Serial position and each of the other two variables were also significant, which were subsumed within a significant three-way interaction between all three variables, $F(5, 95) = 3.72$, $MSE = 0.01$, $p < .01$, which may, speculatively, be described in terms of the phonological similarity effect with single-voice presentation becoming more emphatic across the curve whereas with alternating voices its (generally decreased) magnitude is more constant across the curve (especially across serial positions 3-6). We would not want to attach too much theoretical significance to these interactions involving serial position however: None of the accounts, as far as we are aware, would make particular predictions with regard to the interaction of the two main variables with serial position.

Discussion

Experiment 3 showed that talker variability and phonological similarity interact (i.e., are non-additive) consistent with the view that both effects have the same functional locus. The results are therefore at odds with the predictions of item-decay accounts based on the standard model, at least as exemplified by the phonological loop model (Baddeley, 1986): If the phonological similarity effect is the empirical signature of a passive phonological store, there is no reason to expect it to interact with the talker variability effect which, from this perspective, has been attributed to the articulatory control process (as with the word-length

effect; e.g., Martin et al., 1989). Note that the fact that there was nevertheless a phonological similarity effect regardless of talker variability (albeit a significantly smaller one in the alternating voice conditions) means that the results cannot be accounted for by supposing that the hypothetical phonological store is abandoned under difficult conditions (see e.g., Baddeley, 2000, 2007; Baddeley & Larsen, 2007).

The interaction found between phonological similarity and talker variability is also inconsistent with an item-distinctiveness approach. This approach explains the phonological similarity effect in terms of decreased inter-item distinctiveness. Given that, on this approach, the *greater* inter-item distinctiveness characterizing an alternating voice list should improve performance, the approach must appeal to a mechanism (e.g., attentional-resource depletion; Neath, 2000) other than item-distinctiveness to explain the talker variability effect. Thus, as with the standard model, the non-additive effect found in this experiment is troublesome for this approach.

In contrast, the non-additivity of talker variability and phonological similarity effects is consistent with the perceptual-gestural account. On the view that talker variability impairs the capacity to populate the speech plan with the items in the correct order, any further speech-planning errors—which we have argued elsewhere is primarily responsible for the phonological similarity effect (Jones et al., 2004)—would be expected to have less impact on performance.

Experiment 4

On the perceptual-gestural mismapping account, the assembly of items into a rehearsal cohort (or sequence-output plan in the parlance of the perceptual-gestural view) is a pre-requisite for the effect: The impairment is based on a conflict between by-voice perceptual organization and the tendency to generate a sequence-output plan that mimics the true temporal order of the items. Thus, according to this account, a talker variability

effect should only be produced (or at least be much more pronounced) in short-term memory tasks that engage a sequence-output planning process in support of order reproduction.

In contrast, on the item-decay approach, expression of the talker variability effect should not depend on the requirement for the reproduction of order: Serial recall is vulnerable to talker variability not because of the requirement for *serial* recall per se, but because such variability increases the loss of item-identity, one consequence of which is a difficulty in retrieving the items in correct serial order if the task-instructions happen to demand it. For example, according to the phonological loop model, both item- and order-based short-term verbal memory tasks rely on the phonological store: Henson et al. (2003) contrasted performance on a task emphasizing memory for the identity of the items with that on a task emphasizing order memory and noted that "the observation that both tasks were performed less accurately when probed recalls were [phonologically] confusable provides useful confirmation that they were accessing phonological short-term memory." (p. 1316). Thus, according to any account that conceives of the talker variability effect as reflecting the increased decay of items in phonological short-term memory, a verbal short-term memory task should be vulnerable to talker variability regardless of whether it involves the reproduction of serial order.

Similarly, on the attentional-resource account derived from the feature model, there is no reason to assume that the requirement for *serial* processing is critical for the talker variability effect. This follows from the fact that in the feature model's simulation of the changing-state irrelevant sound effect in serial recall—which appeals to the attentional-resource parameter (Neath, 2000)—the serial nature of the focal task is of no consequence: The value of the attentional parameter is decreased under conditions of changing-state sound regardless of whether the task has a serial component (Neath, 2000).

To date, there exists only scant and equivocal evidence pertaining to whether talker variability effects are found in short-term memory tasks that do not call for order retention (and hence, from our standpoint, gestural-sequencing). For example, Watkins and Watkins (1980) found impaired recall of early list items in a talker variable condition under free recall instructions whereas Martin, Mullennix, Pisoni, and Summers (1987, cited in Goldinger et al., 1991) found that ordered- but not free-recall was vulnerable to a talker variability effect. To complicate matters further, despite the fact that, nominally, free recall requires recall of items in any order, performance of the task is often supported by a rote rehearsal strategy (Bhatarah, Ward & Tan, 2008; Beaman & Jones, 1998; Kahana, 1996). Thus, the free recall task does not hermetically isolate item from order retention processes and cannot therefore speak unequivocally to the issue of whether talker variability affects non-order based tasks. Therefore, in Experiment 4, to examine the role of sequence output-planning in the talker variability effect, we adopt the often-used device of contrasting two tasks in which the presentation conditions and output requirements are identical but the requirement to retain order information differs (e.g., Henson et al., 2003; Jones & Macken, 1993).

We contrast the impact of talker variability on a probed recall task (Waugh & Norman, 1965), which, like serial recall, requires order retention, with that on a missing-item task which requires item, but not order, retention (e.g**.,** Beaman & Jones, 1997; Buschke, 1963; Hughes et al., 2007; Jones & Macken, 1993; Klapp, Marshburn, and Lester, 1983; Macken & Jones, 1995; Murdock, 1993). In the missing-item task, items from a closed set are presented (e.g., permutations of *1-9*) and the task on each trial is to identify which item was left off the list. In this task, then, the emphasis lies with retaining which items were presented so as to be able to identify which one was not: it does not require that the particular order of the items be retained and it is generally assumed that the

task is not performed by recourse to processing their order (e.g., Beaman & Jones, 1997; Macken & Jones, 1995; LeCompte, 1996). The probed recall task involves the same presentation conditions and the same output requirements as the missing-item task (i.e., to produce one item; this latter feature making it a better comparison-task than serial recall) but, at test, one item from the just-presented list is re-presented and the task is to report the item that followed it in the list. In this task, therefore, it is the order of the items that is critical and according to the perceptual-gestural view, performance of this task—like serial recall but unlike the missing-item task—would be supported by a process of converting the incoming sequence into a gestural sequence-output plan. On the basis of the perceptual-gestural mismapping account, therefore, a talker variability effect should be produced in the order-based task (probed recall) but not in the item-based task (missing-item recall). This is because only in the probed recall task can talker variability possibly produce an incompatibility between a sequential perceptual organization defined by voice and a deliberate sequence output-planning process. In contrast, on the item-decay and attentional-resource accounts, talker variability should impair both tasks equally.

Method

Participants

Twenty-six participants from Cardiff University took part in a repeated-measures design in return for course credits. Each participant reported normal or corrected-to-normal vision and normal hearing.

Apparatus and Materials

The apparatus and materials were identical to those used in Experiment 1 except for the following alterations. The to-be-remembered list for each trial consisted of eight digits taken from the 9-item set 1-9 with the item missing from each list, or the item to be probed, chosen randomly for each trial.

Design

The experiment had a repeated-measures design with two factors: List-type (single- or alternating-voice), and Task (missing item or probed recall task). Each participant took part in two blocks of 36 trials. In one block, the task was to identify and recall the missing item whereas in the other block an item (a 'probe') was presented from the to-be-remembered list and participants were required to recall the item that had followed it in the list. Within each block, there were 18 single-voice trials (Single: 9 female, 9 male) and 18 alternating-voice trials (Alt: 9 starting with a female item, 9 with a male item). Within each block, no trial-type was presented more than twice in succession. The order in which the two blocks were undertaken was counterbalanced across participants. There were two practice trials (1 Single, and 1 Alt) preceding each block.

Procedure

The procedure was identical to that of Experiments 1-3 except for the response phase in both blocks. For one block of trials participants were presented with the visual cue "which item was missing?" 50 ms after the offset of the last auditory item. Participants had 10 s in which to indicate—using a keyboard—the digit they thought was missing from the list just presented. For the other block, participants were visually presented with the question "which item followed *x*?" (where *x* represents one of the digits presented in the to-be-remembered list). Participants had 10 s in which to press the numeric key representing the digit they thought followed *x*. As soon as a response was made, or after the 10 s time limit, a 200 ms tone sounded to signal to the participant that the presentation of the first digit of the next to-be-remembered list was imminent. With the inclusion of an optional 5 min break between the two blocks, the experiment lasted approximately 20 min.

Results

Figure 4 shows the percentage of responses in which the missing item (for the missing-item task), and correct digit given the probe (probed recall), were identified in each List-type condition. It is evident that the magnitude of the talker variability effect is much greater for the probed recall task than for the missing-item task. This was confirmed by a repeated-measures ANOVA which showed a main effect of List-type, $F(1, 25) = 55.14$, $MSE = .009$, $p < .01$, and Task, $F(1, 25) = 10.14$, $MSE = .022$, $p < .005$, and, most importantly, a reliable interaction between List-type and Task, $F(1, 25) = 11.40$, $MSE = .009$, $p < .005$, reflecting the fact that the talker variability effect was larger in the probed recall task. However, simple effects analyses (LSD) showed that performance was better with single-voice lists than with alternating-voice lists in both the missing-item ($p = .005$) and probed recall ($p < .001$) tasks.

Discussion

The results of Experiment 4 provide partial support for the perceptual-gestural mismapping account: The effect of talker variability was significantly larger in the probed recall task than in the missing-item task but both tasks were nevertheless impaired. One possible reason why an effect was found in the missing-item task is that even though the retention of order is not an explicit requirement in the missing-item task, serial rehearsal may nevertheless be used to some extent to support performance in this task (see, e.g., Norris, Baddeley, & Page, 2004). However, another possibility is that the particular demands of the missing item task may have produced a 'talker variability effect' that is of a qualitatively distinct form from that found in the probed recall (or serial recall) task. Specifically, the use of two voices may have split the list into two sub-sets within which the search for the missing item had to be conducted. That is, with an alternating-voice list, there is always more than one item missing from within each voice (or each set). It seems plausible that this places an additional burden on identifying the missing item. Thus,

according to this ‘two-sets hypothesis’, an alternating-voice list does not impair missing-item recall by rendering the list perceptually incoherent—as we argue is the case with serial recall and probed recall tasks—but rather because the list is simply not homogeneous in terms of voice. This analysis yields a clear prediction which is tested in Experiment 5: Having two voices convey the list should impair missing-item recall regardless of whether the list is rendered perceptually incoherent.

Experiment 5

This experiment tests the two-sets explanation of the ‘talker variability effect’ in missing-item recall by using, instead of an alternating-voice list, a ‘separated-voices’ condition in which all the items spoken in one voice were spoken before the items spoken in the other voice. Such a condition creates two sets of items but would not render the successive items perceptually incoherent. Thus, if the impairment of missing-item recall in the alternating-voice condition in Experiment 4 was the result of a two-sets problem (and not perceptual incoherence), then the same impairment should be found in a separated-voice condition. This hypothesis gains some credence from a study by Klapp et al. (1983) which found that missing-item recall showed a numerical (but non-significant) impairment when a verbal list was divided into two sets by the insertion of a temporal delay in the middle of the list. Conveniently for analytical purposes, the opposite outcome can be predicted for the probed recall task: In order retention tasks such as probed recall, it is well established that presenting items in sequentially distinct groups—where the groups are separated by, for example, a temporal gap or, indeed, by voice—enhances recall (e.g., Frankish, 1985, 1989; Hitch, Burgess, Towse, & Culpin, 1996; Maybery, Parmentier, & Jones, 2002; Ng & Maybery, 2002; Ryan, 1969). Thus, compared to the single-voice condition, we predicted that separated-voices presentation would facilitate performance in the probed recall task whilst impairing performance in the missing-item task.

Method

Participants

Twenty Cardiff University undergraduates took part in a repeated-measures design in return for course credits. Each participant reported normal hearing and normal or corrected-to-normal vision.

Apparatus and Materials, Design, and Procedure

All these aspects of the methodology were identical to those used in Experiment 4 except for the following alteration: The alternating-voice list condition was replaced with a separated-voices condition whereby one voice conveyed the first four digits of the list and the other voice conveyed the last four digits. For half the trials in this condition, the first four digits were presented in a female voice and the second four digits in a male voice and vice versa for the other half of trials.

Results

Figure 5 shows the percentage of probed and missing items identified correctly for the two list-type conditions. The pattern of results is straightforward: Compared with Experiment 4, the talker variability effect in the context of the missing-item task remains whereas for the probed recall task, separated-voices lists were better recalled than single-voice lists. A repeated measures ANOVA showed that whilst there was no main effect of List-type, $F(1, 19) = 1.66$, $MSE = .005$, $p > .05$, or Task, $F(1,19) = 1.60$, $MSE = .043$, $p > .05$, the critical interaction between List-type and Task was significant, $F(1, 19) = 19.63$, $MSE = .005$, $p < .001$. Simple effects analyses revealed that single- and separated-voices conditions differed from one another in both the missing-item ($p < .05$) and the probed recall ($p < .001$) tasks but in opposite directions.

Discussion

The results of Experiment 5 are consistent with our hypothesis that the mechanism by which talker variability impaired performance of the missing-item task in Experiment 4 is distinct from that involved in tasks calling for order retention: Probed recall—which calls for order retention—was markedly impaired when the voices were alternating (Experiment 4) but was facilitated when the voices were separated into two sequentially distinct sets (Experiment 5) as compared with performance in a single-voice condition. The facilitative effect of grouping on probed recall is consistent with numerous studies showing grouping benefits in order retention tasks (e.g., Frankish, 1985, 1989; Hitch et al., 1996; Maybery et al., 2002; Ryan, 1969). Although the precise mechanism responsible for grouping effects remains a matter of debate, empirically it seems to reflect, at least in large part, a reduction in the probability of transpositions (a type of serial order error in which two neighboring items are switched; cf. Henson, 1998) between items that traverse the boundaries between the 'mini-lists' created by the grouped presentation. In terms of the perceptual-gestural approach, the benefits of grouping may be understood in terms of a particularly good mapping between the organization of the presented material—or the type of gestural organization it promotes—and the temporally-based prosodic habits used in natural language which we assume are co-opted to support serial recall (Jones et al., 2004).

In contrast, the missing-item task was impaired by talker variability regardless of whether the two voices alternated or formed two sequentially distinct sets. This pattern is entirely in line with our hypothesis that the difficulty imposed by talker variability in the missing-item task is caused by having two sets within which to have to search before identifying the missing item and is not restricted therefore to talker-variable lists that are sequentially incoherent. Thus, only when sequence-output planning is likely to be a

dominant component of the task (serial recall and probed recall tasks) does talker variability impair performance by its action of rendering the list perceptually incoherent.

The results of Experiments 4 and 5 pose further problems for item-decay accounts of the talker variability effect. Given that these accounts suppose that the effect results from the greater loss of the integrity of item information, it is far from clear why the missing item task—which would seem to require that the identity of the presented items be retained so as to identify which item was missing—is not susceptible to the 'classical' talker variability effect. Indeed, it could be argued that the missing-item task, according to the item-decay accounts, should be more, not less, susceptible than the probed recall task (or serial recall), in which memory for item content is negligible. However, one possible objection from the perspective of the standard model upon which item-decay accounts are based is that whilst both item- and order-based tasks are assumed to be supported by a labile verbal store (Henson et al., 2003), only order-based tasks (such as the probed recall and serial recall task) involve articulatory rehearsal and only such tasks should, therefore, be susceptible to a talker variability effect. Such an argument would be consistent with the assumption within the standard model that rehearsal is a precondition for the word-length effect, an effect also explained in terms of a race between decay and rehearsal (Baddeley, 2007; Baddeley et al., 1975). In this way, the item-decay accounts as well as the perceptual-gestural mismapping view could account for why only the order-based task showed a talker variability effect. However, a problem with this counterargument is that given that item-based tasks rely on retrieval of decay-prone traces from a phonological store (see Henson et al., 2003), it is far from clear why those traces only need to be refreshed by articulatory rehearsal when their order also needs to be retained.

An attentional-resource based account (e.g., Neath, 2000) also fails to explain why the particular demands of the task has a critical impact on the talker variability effect. It

cannot easily be argued that the missing item task was more quantitatively demanding—and thus required more attentional resources—than the probed recall task and contend on this basis that talker variability could not exert as much damage. This is because in both Experiments 4 and 5, performance levels for the two tasks in the baseline (i.e., single-voice condition) were nearly identical. We suggest that the interactions observed between task and talker variability are, therefore, more parsimoniously ascribed to the qualitatively different requirement for serial order (and hence sequence output-planning) in the probed recall (and serial recall) tasks compared to the missing-item task.

In sum, the results of Experiments 4 and 5 complement those of Experiments 1-3 and, as a set, we suggest that the series provides compelling evidence that the talker variability is a joint product of obligatory auditory perceptual organization and the deliberate attempt to assemble the items into a gestural sequence designed to mimic their true temporal order.

General Discussion

The results of the present series of experiments may be summarized as follows: Experiment 1 showed that the presence of either a single- or alternating-voice lead-in preceding an alternating-voice list accentuates the talker variability effect. According to the perceptual-gestural mismapping account, the lead-ins promoted the perceptual incoherence of the list thereby exacerbating the poor mapping between automatic perceptual order encoding and the need to generate a gestural (articulatory) analogue of the true temporal sequence. Item-decay accounts of the talker variability effect based on the standard model (see Martin et al., 1989; Goldinger et al., 1991) and an attentional-resource account based on the feature model (Neath, 2000) would have predicted that the lead-ins should have reduced rather than augmented the effect.

Experiment 2 showed that the recall of four-voice lists was significantly better than that of a two-voice (i.e., alternating-voice) list despite the fact the four-voice list would be expected to impose a greater burden on item-encoding or attentional resources. The perceptual-gestural mismapping account can readily account for this finding by recourse to the fact that obligatory grouping by acoustic similarity (e.g., provided by a common voice)—and hence a misleading subjective perception of order—would be stronger in the alternating-voice compared to the four-voice condition (Bregman, 1990). The results of this experiment are also not readily accounted for in terms of a deliberate but counterproductive strategy of rehearsing the items by voice (Greene, 1991): It seems unlikely that the four-voice lists would lend themselves to such a strategy and yet a marked impairment was still evident in this condition.

Experiment 3 showed that talker variability reduces the phonological similarity effect in line with the perceptual-gestural view that both effects are located in the gestural-sequencing process. At the same time, this non-additive effect is inconsistent with the item-decay and item-distinctiveness accounts which view the two effects as having different functional loci. Finally, Experiments 4 and 5 together provided further convergent evidence in favor of the perceptual-gestural mismapping account over item-decay and attentional-resource accounts by demonstrating that the (classical) talker variability effect is only found when the task involves order retention and hence gestural rehearsal (probed recall) than when the task only calls for item retention (missing-item task).

Implications for Item-Distinctiveness/Attentional Resource Accounts

The basic talker variability effect would appear to pose a problem for accounts of short-term memory performance couched within an influential class of theory that appeals to the similarity between items in a list (e.g., Brown et al., 2007; Farrell & Lewandwosky, 2002; Nairne, 1990; Neath, 2000): The greater discriminability between items in a talker-

variable list should afford, if anything, better, not poorer, recall. However, we suggested that the feature model, at least, might be able to appeal, instead, to an attentional-resource based account: talker variability might usurp general resources as represented by the attention ('a') parameter included in the model (Neath, 2000).

One difficulty with the feature model's appeal to an attentional parameter (or the concept of attentional resources more generally) is that it has often not been clear, a priori, when an impairment of performance will be attributed to a depletion of attentional resources or to item-distinctiveness mechanisms (see, e.g., Jones & Tremblay, 2000, for a discussion of this issue in relation to the irrelevant sound effect). A key advantage of the present study in this regard is that having talker-variable to-be-remembered items should clearly exert an impact on item-distinctiveness based mechanisms, that is, it should decrease feature overwriting between modality-dependent features of successive items (for details, see Neath, 2000). Thus, to simulate the basic talker variability effect, it would have to be assumed that the depletion of resources resulting from talker variability produces a systematically greater cost than the benefit that the model otherwise predicts should be found with talker-variable lists. Again, it seems that such an assumption can only be made in an ad hoc manner. Moreover, even if we accept this additional assumption, the attentional-resource account fails to explain the more detailed empirical signature of the phenomenon as revealed in the present study. There may of course be ways other than appealing to general attentional resources by which proponents of item-distinctiveness accounts might seek to explain the basic talker variability effect. However, we suspect that the particular pattern of findings revealed in the present study—the role of perceptual organization; the non-additivity of phonological similarity and talker variability; and the particular vulnerability of order-based tasks—may present a considerable challenge to such accounts. We hope that the present results will catalyze efforts to meet such a challenge.

Implications for Standard Model-Based Accounts

Current accounts of the talker variability effect set within the framework of the standard, decay-rehearsal, model of verbal short-term memory (e.g., Atkinson & Shiffrin, 1968; Baddeley, 1986) appeal to the concept of item-decay (Goldinger et al., 1991; Martin et al., 1989; Nygaard et al., 1995). Based on the present results, an item-decay approach to the effect seems no longer tenable, at least in the context of the typical (pure) serial recall setting. However, as noted in the Introduction, although an item-decay approach to the talker variability effect in pure serial recall is logically consistent with the standard model, the item-decay accounts (Goldinger et al., 1991; Martin et al., 1989; Nygaard et al., 1995) were motivated initially by results derived from an arguably atypical serial recall task: each trial comprised 10 unique words on each trial and a free-output procedure was adopted (i.e., items could be output in any order but they had to correspond ultimately to their correct input positions). Moreover, in the talker-variable condition in these studies, each item was spoken in a different voice although it is not clear whether (different) male and female voices nevertheless alternated. It is therefore difficult to assess the extent to which the effects found in this atypical setting are functionally similar to those observed in the present experiments (and in Greene, 1991). For example, presenting a unique set of items on each trial means that the task, unlike pure serial recall, imposes a relatively large burden on item memory. Thus, the talker variability in this setting may affect item, not order, memory even though performance was scored on the basis of a correct correspondence between each item's input and output position (because forgetting and failing to output an item would still have been registered as an error). It is also possible, therefore, that item-decay—based possibly on item-level perceptual factors (voice normalization/incorporation; e.g., Goldinger et al., 1991)—may indeed contribute to the effect found in the atypical serial recall setting. However, such an account would still have to explain why it is that

other tasks calling for item memory (e.g., the missing-item task) seem invulnerable to a 'standard' talker variability effect (present Experiments 4 and 5).

An alternative possibility that would be consistent with the perceptual-gestural mismapping account is that although the atypical serial recall setting (e.g., Martin et al., 1989) places a large burden on item memory, talker variability nevertheless impairs that component of the task that taps into the efficacy of order processing and hence of a sequence output-planning process. Several aspects of previous results are in line with this analysis: First, in both free recall (Watkins & Watkins, 1980) and in the atypical serial recall task (e.g., Martin et al., 1989), talker variability only affects the recall of that part of the list—the early part—that has been found to be supported by serial rehearsal (even in free recall; Beaman & Jones, 1998; Kahana, 1996). Indeed, the serial position curve obtained in the atypical procedure (see, e.g., Martin et al., 1989) resembles a composite of that found in serial recall and free recall: It exhibits both a large, extended, primacy portion (as in serial recall) and a large, extended, recency effect (as in free recall; see Bhatarah et al., 2008). Second, when the rate of presentation is slowed down substantially to beyond 1 item/2000 or 4000ms, the effect of talker variability in the atypical setting reverses to a positive effect (Goldinger et al., 1991). This may be explicable in terms of an articulatory serial rehearsal strategy useful at relatively fast rates (e.g., 1 item/250ms: Goldinger et al., 1991; 1 item/350ms: present study) giving way—particularly given the use of a unique set of semantically-rich items (nouns) for each trial—to a qualitatively different, perhaps semantic-based, strategy at much slower presentation rates, in which the different voices can now be exploited to enhance retention (e.g., by associating each different-voice item with a person's name; Lightfoot, 1989, cited in Goldinger et al., 1991).

Whilst the extant standard model-based accounts of the talker variability effect do not fare well when applied to the effect found in pure serial recall, there may be other

means by which the standard model (particularly the phonological loop model; Baddeley, 1986, 2007) could potentially account for the basic phenomenon. For example, one possibility might be to appeal to computational models of how the phonological loop retains and reproduces serial order, a consideration that has typically been seen as complementary, but secondary, to the core item-based architecture of the underlying functional-level theory (see, e.g., Baddeley, 2007). Thus, talker variability might be seen as disrupting one of the several mechanisms that have been proposed to support serial order in the phonological loop (e.g., a primacy gradient, Page & Norris, 1998; an oscillator-based timing signal; Burgess & Hitch, 1999). Indeed, such an approach has already been adopted in relation to the disruptive impact of irrelevant speech on serial recall (Colle & Welsh, 1976; Salamé & Baddeley, 1982). In principle, therefore, the same approach might be taken in relation to the talker variability effect. Indeed, such an approach would at least allow the standard model to explain our finding that order-based tasks are particularly (or, more arguably, uniquely) susceptible to the effect (Experiments 4 and 5).

However, an approach that appeals to order-mechanisms but nevertheless adheres to the notion of a passive phonological store would still seem to face difficulties with the interaction we observed between phonological similarity and talker variability (Experiment 3). In computational models of the phonological loop, the phonological content of the items—the similarity between which, on this approach, is responsible for the phonological similarity effect—occurs at a stage that is independent of the mechanisms responsible for representing order (Page & Norris, 1998). Thus, if talker variability is associated with the order-storing stage, then it should leave the phonological similarity effect unscathed, contrary to our data. Furthermore, the role played by perceptual organization processes in the effect (Experiments 1-2) would also appear problematic for the approach. The standard model has not, historically, invoked the action of perceptual organization processes (e.g.,

Baddeley, 1986, 2007), presumably because such processes are clearly not specific to verbal input (and hence not uniquely 'phonological'; Bregman, 1990). However, it is possible that the model could be extended so as to include a further module at the front-end of the phonological store to accommodate the impact of perceptual organization processes (e.g., Page & Norris, 2003). Perceptual organization may then be seen as impacting upon a representation of order within the phonological store in some way. However, the difficulty with taking such a step is that as an increasing number of serial recall phenomena are being explained by recourse to general-purpose pre-phonological acoustic (i.e., auditory streaming) processes coupled with articulatory- (or more generally motor-) planning systems, it is becoming less clear what additional explanatory power is gained by postulating a distinct, post-perceptual, memory structure (the phonological store) located in between those perception and action systems (for a convergent view from a neuroscientific perspective, see Buchsbaum & D'Esposito, 2008).

Towards an Embodied View of Short-term Memory

The appeal in our perceptual-gestural framework to general-purpose processes involved in perception and action that are co-opted to meet the demands of a short-memory task (for similar views, see Glenberg, 1997; Reisberg et al., 1984; Wilson & Fox, 2007) resonates with a current shift in cognitive science towards embodying cognition (e.g., Clark, 2006; Pecher & Zwaan, 2005). This shift has emerged as a reaction to the received view of cognition as the action of static, central, and context-free processing and storage structures/resources that are divorced from the so-called "peripheral" processes of perception and action (e.g., Clark, 1999; Hurley, 2001). Instead, an embodied analysis focuses on the dynamic processes involved in goal-directed and coherent engagement with the environment given the constraints and capacities of the organism's sensori-motor apparatus. Thus, in this spirit, one way of fleshing out the gestural component of our

account of short-term memory is to suppose that sequence output-planning (or 'rehearsal') reflects the operation of a motor-action emulator: Recent work on motor control suggests that in order for motor-action to be executed in a fluent manner, a 'forward model' of the action—consisting of both the instructions to the effectors and, importantly, the sensory sequelae of the action—is generated so that the imminent action can be compared with the intended action (e.g., Grush, 2004; Shubotz, 2007). An important feature of these models in the current context is that they can be run without being implemented, that is, they may be run in emulation mode without necessarily resulting in any overt action (e.g., Jordan & Rumelhart, 1992).

Thus, we contend that the tendency to engage in articulatory rehearsal in a verbal serial recall task does not reflect the fact that items are represented in a labile form in a static phonological store (see also Reisberg et al., 1984). Rather, such engagement reflects the fact that the monophonic and hence necessarily sequential nature of speech (or of the emulation of the movements of the vocal tract) endows the individual with an ideal medium for taking a series of largely sequentially-unrelated verbal items and placing them onto a common carrier, that is, a single, relatively more sequentially-coherent, action. In addition, when the to-be-remembered material is auditory-verbal, the action of perceptual organization processes—which are not distinctly phonological—can also come into play to shape performance, as evidenced in the present study (see also Hughes & Jones, 2005; Jones et al., 2004; Nicholls & Jones, 2002): The talker variability effect appears to reflect an impairment in the auditory-motor mapping process (cf. Buchsbaum & D'Esposito, 2008) whereby the process of populating a necessarily abstract motor-output emulator system with specific content is misinformed by the auditory-perceptual organization system (for other examples of the role of pre-phonological perceptual organization processes in

serial short-term memory, see, e.g., Jones, Alford, Bridges, Tremblay & Macken, 1999, Jones et al., 2006; Nicholls & Jones, 2002).

Encouragingly, conclusions from several research programmes other than our own are now converging on an embodied conceptualization of short-term memory. For example, Wilson and Fox (2007) recently found that serial recall of novel sequences of hand gestures (that seemed unlikely to be mediated by their verbal re-labelling) exhibits several of the effects that are, putatively, hallmarks of a specifically verbal short-term memory system, namely, the "phonological" similarity effect, the "articulatory" suppression effect and the "word" length effect. Such results led the authors to suggest that "(r)ather than involving hard-wired and dedicated components, working memory may instead consist of the strategic recruitment of cognitive resources, determined on the fly by the immediate demands of the task" (Wilson & Fox, 2007, p. 473). A trend toward the view that there is no distinct short-term/working memory system is also emerging within the neuroscientific literature. For example, it has been traditional to view the frontal cortex (especially the prefrontal cortex) as the seat of a dedicated short-term/working memory system (e.g., Goldman-Rakic and Leung, 2002; Logie & Della Salla, 2003). However, more recent evidence suggests that activity in these 'working memory' areas may instead reflect non-mnemonic sensory, attentional, and action-related functions involved in an organism's immediate interaction with the environment (for a review, see Postle, 2006; see also Buchsbaum & D'Esposito, 2008). Given such convergence of views, the time seems ripe for a shift in research focus away from delineating the properties of bespoke short-term structures and mechanisms to examining how the capacities and constraints on perceptual and action-planning processes dictate sequential behavior over the short-term.

References

Anstis, S., & Saida S. (1985). Adaptation to auditory streaming of frequency modulated tones. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 257–271.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory, Vol II* (pp. 89–195). New York: Academic Press.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362-365.

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

Baddeley, A. D. (1992). Is working memory working? The fifteenth Bartlett lecture. *Quarterly Journal of Experimental Psychology, 44,* 1–31.

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

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

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology, 36,* 233–252.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *Quarterly Journal of Experimental Psychology, 60,* 497-504.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the

structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575-589.

Beaman, C. P., & Jones, D. M. (1997). The role of serial order in the irrelevant speech effect: Tests of the changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 459-471.

Beauvois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception & Psychophysics, 59,* 81-86.

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

Bregman (1978). Auditory streaming is cumulative. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 380-387.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organisation of sound.* Cambridge, MA: MIT Press.

Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In S. McAdams and E. Bigand (Eds.), *Thinking in sound* (pp. 10–36). Oxford: Oxford University Press.

Bregman, A. S. and Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology, 89,* 244-249.

Bregman, A. S., & Rudnicky, A. I. (1975). Auditory segregation: Stream or streams? *Journal of Experimental Psychology: Human Perception and Performance, 1,* 263-267.

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

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

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

Buchsbaum, B. R.,, & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience, 20,* 762-778.

Buschke, H. (1963). Relative retention in immediate memory determined by the missing scan method. *Nature, 200,* 1129-1130.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 115–127.

Clark, A. (1999). An embodied cognitive science? *Trends in Cognitive Sciences, 3,* 345-351.

Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences, 10,* 370-374.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 17-32.

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

Divin, W., Coyle, K., & James, D. T. T. (2001). The effects of irrelevant speech and articulatory suppression on the serial recall of silently presented lipread lists. *British Journal of Psychology, 92,* 593-616.

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior, 19,* 624–634.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *British Journal of Psychology, 71,* 43-52.

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

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. Journal of Experimental Psychology: Learning, Memory, and *Cognition, 29,* 838-849.

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

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

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20,* 1-55.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Human Perception & Performance, 22,* 1166-1183.

Goldman-Rakic P. S., Leung, H. C. (2002). Functional architecture of the dorsolateral prefrontal cortex in monkeys and humans. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 85–95), Oxford, UK: Oxford University Press.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effect s in recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 152-162.

Greene, R. L. (1991). Serial recall of two-voice lists: Implications for theories of

auditory recency and suffix effects. *Memory & Cognition, 19,* 72-78.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences, 27,* 377-442.

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

Henson, R., Hartley, T., Burgess, N., Hitch, G. & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: a new paradigm and tests of a timing signal hypothesis. *Quarterly Journal of Experimental Psychology, 56A,* 1307-1334.

Hitch, G. J., Burgess, N., Towse, J. N. & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology, 49A,* 140-158.

Houghton, G., & Tipper, S. P. (1996). Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. *Brain and Cognition, 30,* 20-43.

Hughes, R. W., & Jones, D. M. (2005). The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 316-327.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 33,* 1050-1061.

Hurley, S. (2001). Perception and action: Alternative views. *Synthese, 129,* 3-40.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organizational factors in selective attention: The interplay of acoustic

distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 464-473.

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

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology, 60,* 497-504.

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

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, and Cognition, 30,* 656-674.

Jones, D. M., & Tremblay, S. (2000). Interference by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review, 7,* 550–558.

Joos, M. A. (1948). Acoustic phonetics. *Language, 24,* 1-136.

Jordan, M. I., & Rumelhart, D. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16,* 307-354.

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

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Klapp, S. T., Marshburn, E. A., & Lester, P. T. (1983). Short-term memory does not involve working memory of information processing: The demise of a common assumption. *Journal of Experimental Psychology: General, 112,* 240-264.

Lashley, K. S. (1951). The problem of serial order in behaviour. In Jeffress (Ed.),

Cerebral mechanisms in behaviour: The Hixon symposium, (pp. 112-146). New York: Wiley.

LeCompte, D. C. (1996). Irrelevant speech, serial rehearsal and temporal distinctiveness: A new approach to the irrelevant speech effect. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1154-1165.

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

Logie, R. H., Della Salla, S. (2003). Working memory as a mental workspace: Why activated long-term memory is not enough. *Behavioral and Brain Sciences, 26,* 745–746.

Longoni, A. M., Richardson, J. T. E., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory & Cognition, 21*, 11–22.

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner Ear: Single or double agency? *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 436-448.

Macken, W. J., & Jones, D. M. (2003). Reification of phonological storage. *Quarterly Journal of Experimental Psychology, 56A*, 1279-1288.

Magnuson, J. S., & Nusbaum, H. C. (2007). Acoustic differences, listener expectations, and the perceptual accommodation of talker variability. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 391-409.

Martin, C. S. Mullennix, J. W., Pisoni, D. B., & Summers, W.V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 676-684.

Maybery, M. T., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal

memory. *Journal of Memory & Language, 47*, 360-385.

Murdock, B. B., Jr. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review, 100*, 183-203.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology, 78*, 679-684.

Murray, A., & Jones, D. M. (2002). Articulatory complexity at item boundaries in serial recall: The case of Welsh and English digit span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 594–598.

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

Nairne, J. S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology, 53*, 53-81.

Neath, I. (2000). Modelling the effect of irrelevant speech on memory. *Psychonomic Bulletin and Review, 7*, 403-423.

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

Nicholls, A. P., & Jones, D. M. (2002). Capturing the suffix: Cognitive streaming in immediate serial recall. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28*, 12-28.

Norris D., Baddeley A. D., Page, M. P. A. (2004). Retroactive effects of irrelevant speech on serial recall from short-term memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30, 1093-1105.

Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1995). Effects of stimulus

variability on perception and representation of spoken words in memory. *Perception & Psychophysics, 57,* 989-1001.

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

Page, M. P A. & Norris, D. G. (2003). The irrelevant sound effect: What needs modelling and a tentative model. *Quarterly Journal of Experimental Psychology, 56A,* 1289-1300.

Pecher, D., & Zwaan, R. A. (2005). *Grounding cognition. The role of perception and action in memory, language, and thinking.* Cambridge University Press.

Poirier, M., & Saint-Aubin, J. (1996). Immediate serial recall, word frequency, item identity and item position. *Canadian Journal of Experimental Psychology, 50,* 408–412.

Postle, B. R. (2006). Working Memory as an emergent property of the mind and brain. *Neuroscience, 139,* 23-38.

Reisberg, D., Rappaport, I., & O'Shaughnessy, M. (1984). Limits to working memory: The digit digit-span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 203–221.

Repovs, G., & Baddeley, A. D. (2006). Multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience Special Issue, 139,* 5-21.

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

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended

speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150–164.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11,* 211-218.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 419–425.

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

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1,* 86-98.

van Noorden, L. P. A. S. (1975). Temporal coherence in the perception of tone sequences. Unpublished doctoral thesis, Eindhoven University of Technology.

Warren, R. M. (1999). *Auditory Perception: A New Analysis and Synthesis.* New York: Cambridge University Press.

Warren R. M., Obusek, C.J., Farmer, R. M., & Warren, R. P. (1969). Auditory sequence: Confusion of patterns other than speech or music. *Science, 164,* 586-587.

Watkins, O. C., & Watkins, M. J. (1980). Echoic memory and voice quality: Recency recall is not enhanced by varying presentation voice. *Memory & Cognition, 8,* 26-30.

Waugh, N. C. & Norman, D. A. (1965). Primary memory. *Psychological Review,* 72, 89-104.

Wilson, M. & Fox, G. (2007). Working memory for language is not special: Evidence for an articulatory loop for novel stimuli. *Psychonomic Bulletin & Review 14,* 470-473.

Woodward, A. J., Macken, W. J., & Jones, D. M. (2008). Linguistic Familiarity in

Short-Term Memory: A Role for (Co-)Articulatory Fluency? *Journal of Memory and Language, 58,* 48-65.

Footnotes

1. The term ‘talker variability effect’ has been used in several domains (e.g., in the context of long-term recognition tasks; see Goldinger, 1996). However, in the present article, when using this term, we are referring specifically to the effect in the context of serial recall.

Author note

Robert W. Hughes, John E. Marsh, and Dylan M. Jones, School of Psychology, Cardiff University, UK. Dylan Jones is also adjunct professor at the Department of Psychology, University of Western Australia. The research reported in this article received financial support from the United Kingdom's Economic and Social Research Council in the form of a grant awarded to Dylan Jones, Robert Hughes, and Bill Macken. The results of some of the studies comprising the current article have been presented at the Annual Meeting of the *British Psychological Society*, Cardiff, April, 2006; the 5th Annual *Auditory Perception Cognition and Action Meeting* (APCAM), Houston, Texas, November, 2006; the Meeting of the *Experimental Psychology Society*, Cardiff, UK, April, 2007, and at the 48th Annual Meeting of the *Psychonomic Society*, Long Beach, California, November, 2007. Thanks are due to Bill Macken for several useful discussions concerning the studies reported here. Correspondence can be addressed to Robert Hughes at the School of Psychology, Cardiff University, PO Box 901, Cardiff CF10 3AT, United Kingdom. Email may be sent to HughesRW@cardiff.ac.uk.

Table captions

Table 1. A schematic representation of the six conditions contrasted in Experiment 1. Single = Single voice; Alt = Alternating voices. For conditions 3-6, the first part of each condition-name refers to the voice presentation-format of the lead-in whilst the second refers to that for the to-be-remembered list.

Figure captions

Figure 1. Mean percentage of items correctly recalled at each serial position in the six conditions of Experiment 1 (see Table 1 for an illustration of the six conditions).

Figure 2. Mean percentage of items correctly recalled at each serial position in the Single-, Alt- (Alternating-voice), and Four-voice conditions in Experiment 2.

Figure 3. Mean percentage of items correctly recalled in the single- and alternating-voice (Alt) conditions for phonologically dissimilar (Diss) and phonologically similar (Sim) lists in Experiment 3.

Figure 4. Mean percentage of items correctly recalled in the single- and alternating-voice conditions in the Missing Item and Probed recall tasks in Experiment 4.

Figure 5. Mean percentage of items correctly recalled in the single- and separated-voices conditions in the Missing Item and Probed recall tasks in Experiment 5.

Table 1

Condition	Voice	Lead-in	To-be-remembered list
1. Single	Female (or male) voice:		6 5 2 7 1 4 8 3
2. Alt	Female (or male) voice:		6 2 1 8
	Male (or female) voice:		5 7 4 3
3. Single-Single	Female (or male) voice:	8 7 6 5 4 3 2 1	6 5 2 7 1 4 8 3
4. Alt-Alt	Female (or male) voice:	8 6 4 2	6 2 1 8
	Male (or female) voice:	7 5 3 1	5 7 4 3
5. Alt-Single	Female (or male) voice:	8 6 4 2	6 5 2 7 1 4 8 3
	Male (or female) voice:	7 5 3 1	
6. Single-Alt	Female (or male) voice:	8 7 6 5 4 3 2 1	6 2 1 8
	Male (or female) voice:		5 7 4 3

Figure 1

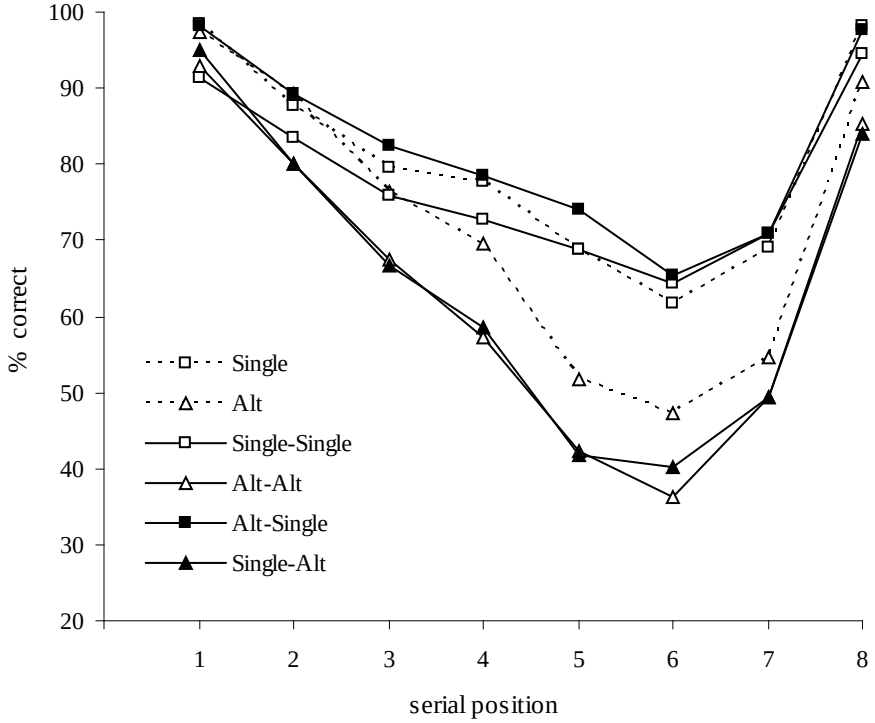

100
90
80
70
60
50
40
30
20
% correct
Single
Alt
Single-Single
Alt-Alt
Alt-Single
Single-Alt
1
2
3
4
5
6
7
8
serial position

Figure 2

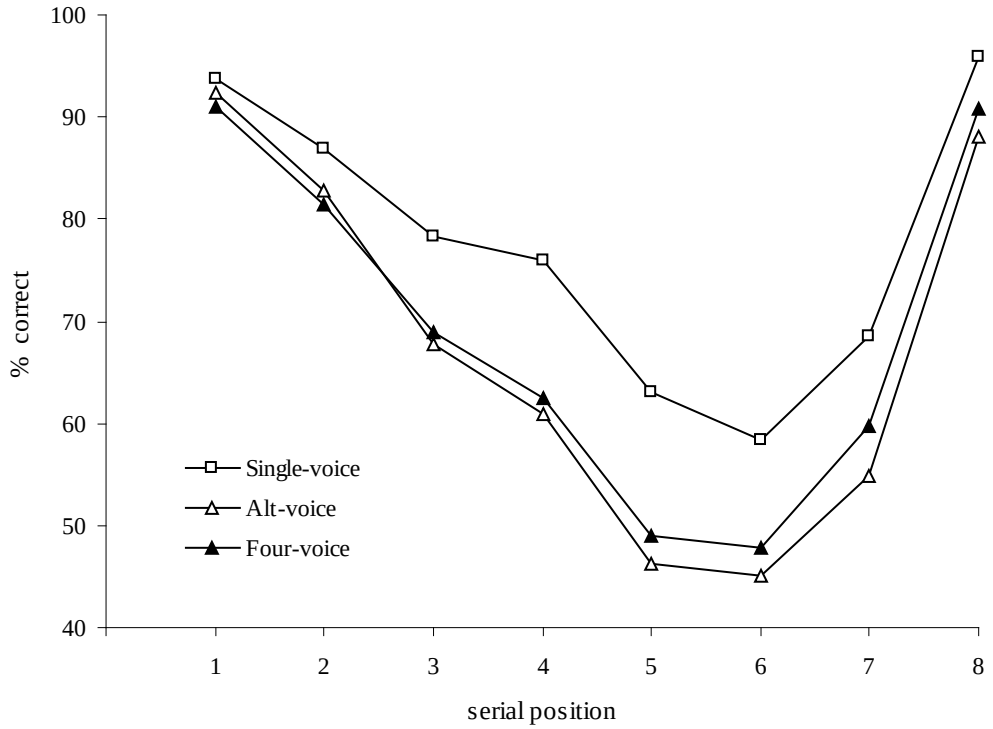

100
90
80
70
60
50
40
% correct
Single-voice
Alt-voice
Four-voice
1
2
3
4
5
6
7
8
serial position

Figure 3

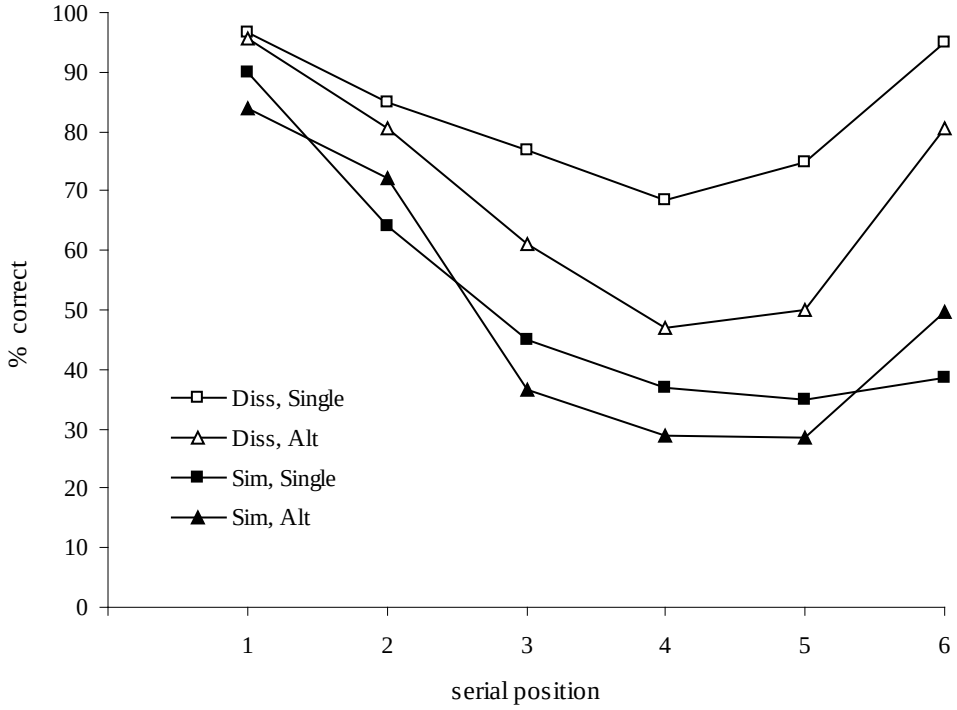

100
90
80
70
60
50
40
30
20
10
0
% correct
1
2
3
4
5
6
serial position
Diss, Single
Diss, Alt
Sim, Single
Sim, Alt

Figure 4

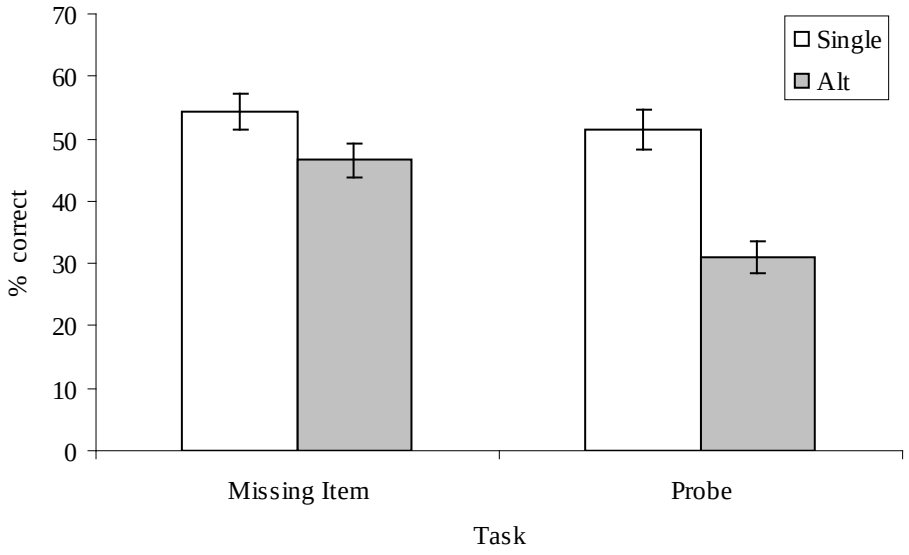
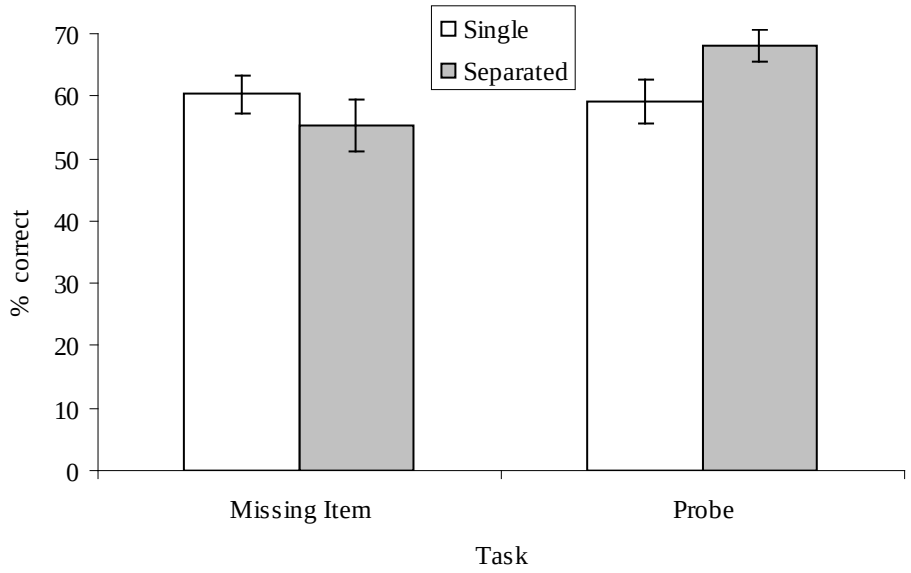

70
60
50
40
30
20
10
0
% correct
Single
Alt
Missing Item
Probe
Task

Figure 5


Single
Separated
70
60
50
40
30
20
10
0
% correct
Missing Item
Probe
Task