Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Memory and Language 54 (2006) 265–281

Journal of Memory and Language

www.elsevier.com/locate/jml

Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory

Dylan M. Jones, Robert W. Hughes *, William J. Macken

School of Psychology, Cardiff University, Park Place, Cardiff CF10 3AT, UK

Received 16 August 2005; revision received 26 October 2005
Available online 20 December 2005

Abstract

Three experiments examined whether the survival of the phonological similarity effect (PSE) under articulatory suppression for auditory but not visual to-be-serially recalled lists is a perceptual effect rather than an effect arising from the action of a bespoke phonological store. Using a list of 5 auditory items, a list length at which the expression of phonological storage should, ostensibly, be strong, the PSE under suppression was removed at recency by a suffix (Experiment 1) and removed throughout by a suffix combined with a prefix (Experiment 2). Finally, the PSE under suppression could be restored simply by decreasing the acoustic similarity between the prefix-and-suffix and the to-be-remembered list (Experiment 3). The results favour a perceptual-gestural view over a dedicated-system view of short-term 'memory.'

Keywords: Short-term memory; Serial recall; Auditory perceptual organization; Phonological store; Phonological similarity

Introduction

Most models of cognition view perceptual processes as being distinct from those responsible for the short-term retention of information. That is, short-term memory performance is usually explained by invoking a bespoke short-term memory system that is intrinsically post-perceptual and functionally divorced from 'earlier,' pre-categorical, perceptual processes (e.g., Atkinson & Shiffrin, 1968; Baddeley, 1986; Baddeley & Hitch, 1974; Broadbent, 1958; Cowan, 1995; but see Crowder, 1993). Indeed, the boundary between pre-categorical perceptual processes and post-categorical mnemonic processes is seen by some as "a fundamental division in cognitive architecture" (Pashler, 1998, p.33; but see Houghton, Macken, & Jones, 2003). However, there is now a growing body of evidence suggesting that many effects traditionally classed as short-term memory phenomena can be better and more parsimoniously understood by recourse to more "peripheral" processes of auditory perceptual organization and gestural skills (e.g., speech) that are co-opted opportunistically to meet the demands of the short-term 'memory' task (e.g., Hughes & Jones, 2005; Jones, Macken, & Harries, 1997; Jones, Macken, & Nicholls, 2004; Macken & Jones, 2003; Nicholls & Jones, 2002a, 2002b; Woodward, Macken, & Jones, 2005).

* Corresponding author. Fax: +44 2920 874858.
E-mail address: HughesRW@cardiff.ac.uk (R.W. Hughes).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2005.10.006

The present article is concerned with revealing the perceptual contribution, more specifically from the auditory system, to phenomena hitherto ascribed to distinctly mnemonic processes. Specifically, we examine an interaction among three factors that has played a pivotal role in the shaping and sustaining of the idea of a short-term phonological store, a post-perceptual component of the cognitive architecture designed specifically for the temporary storage of abstract representations of verbal events (cf. the working memory model; Baddeley, 1986; Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998). The interaction in question is that the phonological similarity effect (PSE; Conrad, 1964) survives under articulatory suppression when the mode of presentation is auditory but not when presentation is visual (Baddeley, Lewis, & Vallar, 1984; Jones et al., 2004; Murray, 1968; Peterson & Johnson, 1971). The proposition to be examined here is that this key interaction is due not to phonological storage but due entirely to acoustically based auditory perceptual organization. Although there is already some evidence supporting this perceptual organization account (Jones et al., 2004), the present article addresses the critical question of whether the account generalizes to a situation in which, ostensibly, the expression of phonological storage should be particularly strong, namely, when the to-be-remembered lists are relatively short (see, e.g., Baddeley, 2000b; Baddeley & Larsen, 2003; Salamé & Baddeley, 1986). We turn first to describe the currently dominant, phonological store-based, account of the interaction between phonological similarity, modality and articulatory suppression.

Phonological storage account

The phonological store retains representations of verbal events that are subject to loss due to decay unless revivified in some way. Phonological representations can be revived by the action of an *articulatory control process* which re-cycles representations through the store, the phenomenal realization of which is sub-vocal rote rehearsal (e.g., Baddeley, 1986, 2000a). The empirical signature of the action of the phonological store is the PSE: items that sound alike (e.g., "*b d g t c*") are serially recalled more poorly than dissimilar sounding items ("*f q r h y*"; Conrad, 1964; Conrad & Hull, 1964), even when presented visually. The phonological storage view is that this effect arises due to confusions between similar phonological representations when retrieving items from the phonological store, although the precise mechanism has not been detailed.

A further key assumption of the phonological store construct is that, for auditory input, entry into the phonological store is automatic and obligatory whereas for visual input there is a more deliberate, and hence optional, grapheme-to-phoneme conversion process involving the articulatory control process. This process not only allows for visual-verbal input to be converted into phonological form, it is also, as noted earlier, the vehicle by which the decay-prone representations within the store (regardless of their modality of origin) are revivified. Note, therefore, that the PSE found with auditory presentation and that found with visual presentation are both explained by recourse to the action of the same short-term phonological store. Indeed, '… the two effects are of roughly comparable magnitude and show an essentially similar pattern of performance across a range of manipulations' (Baddeley et al., 1984, p. 242).

An important logical consequence of the fact that the phonological store does not discriminate between input from external acoustic input and input from internal speech is that the representations with which the store deals—contrary to the interpretation of Wilson (2001)—can neither be acoustic nor articulatory but must, rather, be post-categorical, "central" representations that are functionally remote from more peripheral perceptual or motoric systems. Indeed, the use of the term *phonological* seems to have been deliberately adopted in favor of the terms *acoustic* or *articulatory* (see, e.g., Baddeley, 2002) to indicate the abstract nature of the phonological store's unit of currency. Thus, the phonological store's representations are, by definition, post-categorical, canonical representations of the sounds of a language. They are static linguistic representations that bear very little resemblance to the dynamic and context-sensitive pattern of acoustic energy conveyed by speech, the pattern of light conveyed by the written word, or the movement of the articulators during (covert or overt) speech production (see, e.g., Porter, 1987). According to the working memory model, then, the PSE is an impairment of short-term memory resulting from a similarity between abstract representations of linguistic units regardless of whether these linguistic representations have been abstracted from actual speech sounds or from the written word.

That auditory and visual input gain access to the phonological store by different means provides an explanation for the intricate interplay of three factors that is the central interest of the present article: modality of list presentation, articulatory suppression, and phonological similarity. Articulatory suppression serves to block entry of visual-verbal material into the phonological store whereas auditory-verbal material, in contrast, enjoys direct access and its recall is therefore relatively unimpaired by suppression. The key empirical signature of the action of the phonological store—the PSE—should therefore remain evident during articulatory suppression, but only with auditory material, which enjoys the direct access route, and not with visual material. And indeed, at a macroscopic level, the effect of the combination of these factors is entirely in line with the predictions (Baddeley et al., 1984; Jones et al., 2004; Murray, 1968; Peterson & Johnson, 1971) and has been taken as strong support for the phonological store construct and the working memory model generally (see, e.g., Wilson, 2001). However, a more detailed analysis reveals that this interaction may have arisen from sources distinctly different from those originally proposed, namely, from the action of auditory perceptual, not phonological, factors. That is, the interaction seems to depend on properties that emerge from the acoustic (and hence physical) similarity of the to-be-remembered items, not their phonological (and hence abstract) similarity. We turn now to discuss the evidence for this alternative viewpoint.

Perceptual organization account

Doubts regarding the phonological origins of the interaction between similarity, modality, and suppression arise specifically from detailed consideration of serial position data. Despite the interaction having been demonstrated several times and despite the amount of theoretical weight placed upon it, only rarely have serial position data been presented (Jones et al., 2004; Murray, 1968). It turns out that the precise pattern of interaction in relation to serial position proves problematic from the phonological storage view. Under suppression, the survival of the PSE with auditory presentation is due largely to the survival of performance for the last one or two phonologically dissimilar items (i.e., the *recency* portion of the serial position curve; Jones et al., 2004).

That the effect is localized in recency suggests that the interaction in question is best understood by recourse to the contribution of auditory perceptual factors to short-term serial recall. Better recall of the last one or two items in an auditory list compared to the recall of the same items with visual presentation is one of the canonical features of serial recall, this difference being almost entirely responsible for the superior recall of auditory over visual lists (the *modality effect*; for a discussion, see Penney, 1989). This particularly pronounced recency found with auditory lists (hereafter *auditory recency*) has been explained in a number of ways, but none of these explanations invokes explicitly phonological processes; instead, explanations embodying acoustic factors have been deployed (for a review, see Nicholls & Jones, 2002a). Empirically, the acoustic rather than phonological character of auditory recency can be demonstrated by adding a redundant end-of-list suffix to auditory lists, a procedural step that sharply reduces auditory recency (and hence diminishes the modality effect; Crowder & Morton, 1969). A suffix that shares the acoustic character of the list is typically the most disruptive. Its effect is sharply attenuated if the suffix differs from the list in terms of, for example, voice (e.g., Greenberg & Engle, 1983; Greene, 1991), spatial location (e.g., Frick, 1988; Morton, Crowder, & Prussin, 1971) or rhythm (e.g., Crowder, 1971; Frankish & Turner, 1984). In each of these cases, the phonological identity of the suffix remained fixed, only its acoustic character was changed (for a general discussion of auditory grouping effects, see Frankish, 1985, 1989).

Auditory recency and the action of a suffix can be explained in terms of automatic auditory perceptual organization. There is a rich body of work demonstrating that the automatic encoding of order in auditory sequences is particularly strong for stimuli at or near the boundaries of a sequence at least when the stimuli themselves are acoustically distinct from one another[1] such as would be the case for an auditorily presented "phonologically" dissimilar list (e.g., "*h j q r*..." compared with "*b c g t*..."; see Bregman, 1990; Nicholls & Jones, 2002a; see also Frankish, 1989). In this light, we may view the to-be-remembered list as a temporally extended auditory 'object' obeying rules of perceptual organization—generically referred to as 'auditory streaming'—possessing boundaries whose salience promotes the preservation of order (see Bregman, 1990; Warren, 1982). We suggest that the participant—particularly under conditions of articulatory suppression—draws upon the products of auditory perceptual processing to recover the order of items at the end boundary of the to-be-remembered list. That order encoding is particularly strong for stimuli that are acoustically distinct from one another accounts for the localization of the survival of a PSE at recency under conditions of articulatory suppression (Jones et al., 2004).

This perceptual organization account of the survival of the PSE under suppression with auditory presentation predicts that a suffix should abolish that similarity effect. That is, the presence of a suffix means that the last items in the to-be-remembered list are now remote from the boundary, thereby rendering less effective the strategy of exploiting acoustic-based order encoding for those items. A study by Jones et al. (2004) confirmed this prediction: the survival of the similarity effect under suppression is indeed abolished simply by adding a suffix. On the basis of this and other findings, Jones et al. (2004) argued that the notion of phonological coding is an encumbrance unnecessary to theorizing about short-term verbal memory. To explain the critical interaction, all that needs to be supposed is that factors related to the acoustic characteristics of the sequence sustain auditory recency under suppression and that the PSE observed when participants are free to rehearse (with both visual and auditory input) is located in the articulatory rehearsal process itself, not within a separable,

[1] It should be noted that these acoustic changes must represent variations on a common ground. Thus, the sequence "*h j q r*..." spoken in the same voice meets this condition *par excellence* (see Jones & Macken, 1995a; Jones, Saint-Aubin, & Tremblay, 1999; Warren & Obuzek, 1972).

passive, store that is fed by that rehearsal process (see General discussion).

Present study

A potential stumbling block for the perceptual organization account of the interaction between similarity, modality, and suppression is that such an explanation may only apply to settings in which the expression of phonological storage may be weak. To elaborate, Jones et al. (2004) used a list length of 7 and it has been suggested that at this list length—at least in conjunction with the presence of a suffix and/or a requirement for articulatory suppression—the task becomes difficult enough to lead participants to abandon the use of the phonological store and begin to rely on other, non-phonological, coding mechanisms. Indeed, this argument has often been used in other contexts to explain away data that are otherwise troubling for the phonological store construct (e.g., Baddeley, 1966a, 1966b, 2000a, 2000b; Baddeley & Larsen, 2003; Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987, see also Neath, Bireta, & Surprenant, 2003) and has the potential to undermine the degree to which the perceptual organization account provides a comprehensive explanation of the critical interaction in question (but see Discussion of Experiment 2 for a critique of this ''store abandonment'' argument).

The importance to theory development of precisely why the interaction between similarity, modality, and suppression arises should not be underestimated. Baddeley clearly considers the interaction pivotal to the success of his working memory model: ''The particular pattern of results obtained was crucial to separating the two components of the articulatory loop, the phonological store and the articulatory control process. Had the results not worked out in this way, it would have been necessary to modify the model quite seriously'' (Baddeley, 1986, p. 257). If, therefore, the interaction is indeed mediated by acoustic, not phonological, coding (Jones et al., 2004), the veracity of the working memory model is called into question. By the same token, however, the perceptual organization account would be compromised if it turns out to be of limited generality. What is required, therefore, is a study that sets the phonological storage and perceptual organization accounts in opposition within a setting in which, hypothetically, the action of the phonological store is likely to be particularly dominant, namely, with short to-be-remembered lists.

From the standpoint of the perceptual organization account of the survival of the PSE with auditory items under suppression, the same result is predicted with short lists (5 items) as was found previously with longer lists (7 items; Jones et al., 2004): the survival of the PSE will be confined to recency (i.e., it is a reflection of acoustic, rather than phonological, characteristics of the list). The phonological storage view predicts that the similarity effect for auditory lists under suppression at this shorter list length will extend beyond recency, signifying thereby the general action of the phonological store, and not just the products of acoustic coding.

Experiment 1

In the first experiment, we examine whether the survival of the PSE under suppression with auditory lists is localized in recency with 5 item lists. Moreover, we examined whether this residual similarity effect is abolished by a redundant end-of-list suffix (''*go''*) presented in the same voice as the list items. The removal of the PSE by a suffix with 5-item lists would provide strong support for the perceptual organization account whilst seriously undermining the phonological storage account. Throughout the current series, we used an order reconstruction procedure (e.g., Healy, 1974) whereby following the auditory presentation of the last item, the letters presented on that trial were re-presented simultaneously on a screen (in a fixed arrangement for each trial) and the participants' task was to serially recall the items by clicking each item in the correct serial order. Although the response mode in Jones et al. (2004) was written, there are no a priori grounds for supposing that this procedural difference between that study and this should compromise attempts to compare the results of the two studies. Indeed, the order reconstruction procedure may in fact be the more appropriate or sensitive tool for examining the PSE given that it is an effect mediated by an increased propensity for order errors (transpositions) between items within the same list when those items are ''phonologically'' similar, rather than by item-based errors such as omissions or intrusions (e.g., Bjork & Healy, 1974; Poirier & Saint-Aubin, 1996) which are possible to commit with the written response procedure but not with the order reconstruction procedure. In other words, when examining the PSE, item-based errors are simply an additional source of noise that is prevented here by the use of order reconstruction.

Method

Participants

Twenty-four undergraduate students from the School of Psychology at Cardiff University were awarded course credits for their participation. All were native English speakers reporting normal hearing and normal or corrected-to-normal vision.

Apparatus and materials

The to-be-remembered sequences consisted of 5 to-be-remembered letters, which were selected without replacement either from a closed set of 5 dissimilar consonants (*r, k, l, h, q*) or from a closed set of 5 phonologically similar consonants (*c, b, d, g, t*). These letters, as well as the word "*go*," were recorded in a male voice at an approximately even pitch. Tokens were digitized with a 16-bit resolution, at a sampling rate of 48 kHz using the *SoundForge 7.0* software and edited digitally so that each had a duration of 250 ms. The items (including the suffix if present; see Design below) were presented one-by-one with an inter-stimulus interval (ISI; offset to onset) of 750 ms and at a sound level of approximately 65 dB(A) using the *Psyscope* software (Cohen, MacWhinney, Flatt, & Provost, 1993) running on a *Macintosh Performa.* The response mode for all experiments reported in this article was manual order reconstruction: participants were to reconstruct the order of the items by mouse clicking on buttons labelled with the to-be-remembered letters. The order reconstruction program was written and presented using *Visual Basic* and ran on a PC placed next to the *Macintosh Performa.*

Design

A repeated-measures design was employed with Similarity, Suffix (present or absent), Suppression and Serial position as factors. The experiment consisted of 2 blocks of trials. Both blocks consisted of 48 lists: 24 phonologically similar and 24 phonologically dissimilar lists. Twelve of the phonologically similar lists and 12 of the phonologically dissimilar lists were followed by the suffix "*go*." The onset of the suffix occurred 750 ms following the offset of the final to-be-remembered item and thus the suffix adhered to the rhythm of the to-be-remembered items. Note that regardless of whether a suffix was presented, the recall cue occurred 50 ms following the offset of the final to-be-remembered item. This meant that the presence of a suffix was not confounded with the amount of delay between the offset of the last item and the point at which participants were free to start responding using the order reconstruction display. The four types of trial were presented across the 48-trial block in a pseudo-random order with the constraint than no condition was presented more than twice in succession. One 48-trial block was undertaken under articulatory suppression whilst the other was undertaken without. Whether participants engaged in suppression during the first block or the second was counterbalanced.

Procedure

All participants were tested individually in a sound-proof booth. Each participant first read standard instructions which informed them of what the serial recall task involved and instructed them to ignore the word "*go*" when heard at the end of some of the lists. Participants were informed also that the trials would be presented at a pre-set pace: 50 ms following the offset of the last item, the screen flashed from white to black for 150 ms which signalled to the participant that they should begin to reconstruct the order of the items. From the offset of the screen flashing, there were 15 s before the presentation of the first item of the next to-be-remembered list. During this response period, participants were to click on each of the 5 letters in the order they thought they had been presented. Once a letter was clicked the button changed colour and that button could not be clicked again during the trial. Having clicked on the final item, participants clicked on a 'validate input' button and then clicked off the order reconstruction display (so that the display could not be used as a way of rehearsing the lists; participant clicked back onto the display following the subsequent recall cue).

Thirteen seconds into the 15 s recall interval, a 500 ms tone was presented over the headphones to signal to the participant that the presentation of the first item of the next sequence was imminent (and to start suppressing during 'with suppression' blocks; see below). Participants were told before one of the blocks to engage in articulatory suppression. Specifically, once they heard the tone played over the headphones—which occurred 2 s before the onset of the first to-be-remembered item of each list—they were to start whispering the sequence *"eight, nine, ten"* over and over into a microphone at a rate of approximately three words per second until the recall cue (the screen flashing). The experimenter coached each participant in the correct rate and loudness for the articulatory suppression. Moreover, with the participant's permission, compliance with this instruction was ensured by the experimenter via a sound link. There were two practice trials preceding each block; one similar list without a suffix and one dissimilar list with a suffix. There were therefore 100 trials in all. Including an optional 10 min break between the 2 blocks, the experiment lasted approximately 70 min.

Results

The raw data for all experiments in this article were scored according to a strict serial order criterion: an item was only scored as correct if it was recalled in its correct absolute serial position. The data are portrayed first in general form: Fig. 1 shows the percentage of correctly recalled items in each of the eight conditions (2 [suppression] by 2 [similarity] by 2 [suffix]) without respect to serial position. There is evidence of a PSE in all pairs of conditions although the effect appears to be attenuated in the presence of a suffix regardless of suppression condition. To scrutinize the data further, Fig. 2 shows the mean percentage of items correctly recalled across the five serial positions in the similar and dissimilar conditions, with and without a suffix, in the 'no suppression' (panel A) and 'with suppression' (panel B) conditions. It is evident that without suppression the similarity effect

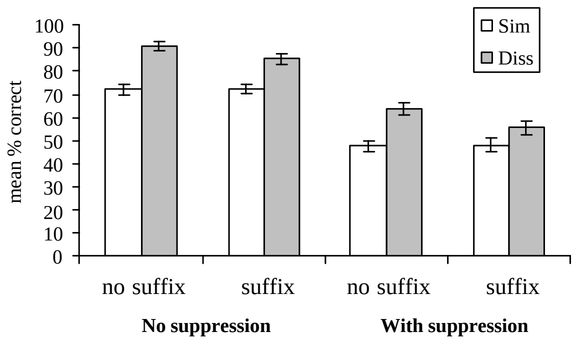

Fig. 1. Mean percentage of phonologically similar (Sim) and dissimilar (Diss) items correctly recalled (without respect to serial position) with and without a suffix in the 'no suppression' and 'with suppression' conditions in Experiment 1. Error bars represent the standard error of the mean.

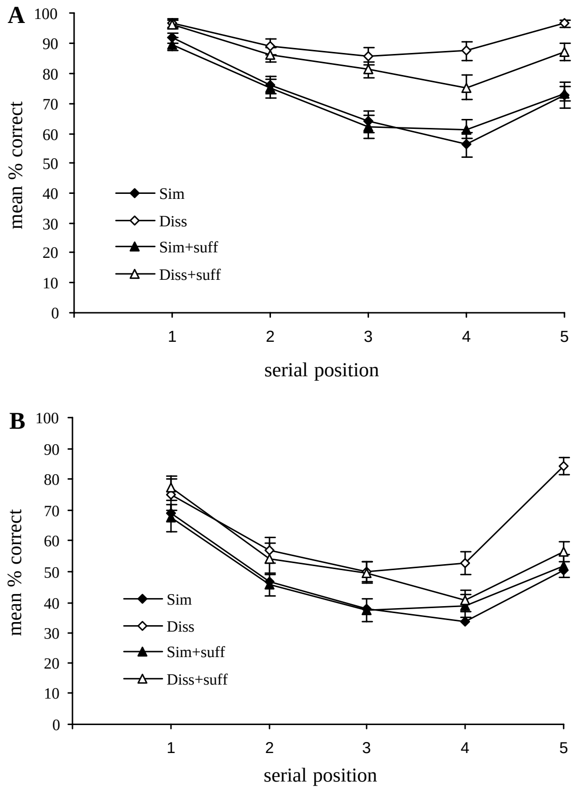

Fig. 2. Mean percentage of phonologically similar (Sim) and dissimilar (Diss) items correctly recalled across serial positions with and without a suffix (suff) in the 'no suppression' (A) and 'with suppression' (B) conditions in Experiment 1. Error bars represent the standard error of the mean.

was present throughout the curve regardless of the suffix, although the suffix slightly attenuated the similarity effect at recency. The more critical data are those shown in panel B (with suppression). In the absence of a suffix there is, as predicted by the phonological storage account, evidence of a PSE throughout the list. Moreover, although a suffix abolished the similarity effect at the last two serial positions, it had no impact on the similarity effect evident at earlier positions.

Given that the design involved a relatively large number of factors and because we are concerned mainly with the precise pattern of performance within the 'with suppression' conditions, separate repeated-measures ANOVA were conducted on the data from the 'no suppression' (Fig. 2A) and 'with suppression' (Fig. 2B) conditions. The ANOVA on the 'no suppression' data revealed main effects of Similarity, $F(1, 23) = 116.08$, $MSE = 3.77$, $p < .001$, and Serial position, $F(4, 92) = 44.71$, $MSE = 2.57$, $p < .001$. There were also trends, albeit non-significant ones, for a main effect of Suffix, $F(1, 23) = 3.59$, $MSE = 3.90$, $p = .071$, and an interaction between Suffix and Similarity, $F(1, 23) = 3.80$, $MSE = 4.05$, $p = .063$, which were subsumed under a reliable three-way Similarity × Suffix × Serial position interaction, $F(4, 92) = 4.18$, $MSE = 1.20$, $p < .01$. This three-way interaction, as well as the trends for a main effect of Suffix and a Suffix × Similarity interaction, can be understood by recourse to the impact of the suffix at recency for dissimilar but not similar items. The only other reliable effect was a Similarity × Serial position interaction, $F(4, 92) = 9.78$, $MSE = 1.7$, $p < .001$, reflecting the fact that, generally, the similarity effect was greater at later than at earlier serial positions.

Turning to the analysis more critical to the central interest of the present article, an ANOVA on the data portrayed in Fig. 2B showed main effects of Similarity, $F(1, 23) = 21.02$, $MSE = 11.33$, $p < .001$, Serial position, $F(4, 92) = 53.78$, $MSE = 4.12$, $p < .001$, and Suffix, $F(4, 92) = 9.59$, $MSE = 9.59$, $p < .01$. All possible interactions were significant (to at least the $p < .05$ level) including the three-way interaction, $F(4, 92) = 10.10$, $MSE = 1.66$, $p < .001$, reflecting the fact that the suffix abolished the similarity effect at the final two serial positions.

Discussion

The results of Experiment 1 are unambiguous in showing that with relatively short auditory lists (5 items), unlike the case with relatively long lists (7 items; Jones et al., 2004), the survival of the similarity effect under articulatory suppression is not restricted to recency but is apparent throughout the serial position curve. A suffix abolishes the PSE in recency but leaves intact the PSE evident in primacy. For recency at least, the logic we applied to the action of the suffix for 7-item lists in Jones et al. (2004) still seems appropriate: the survival of auditory recency for dissimilar but not similar items is due to the action of acoustically driven perceptual organization processes and the suffix disrupts this organization by reducing the distinctiveness of the order of the terminal list items. However, on the face of it, the fact that there was a PSE for pre-recency items suggests that the perceptual organization account (Jones et al., 2004) does not provide a comprehensive explanation of the critical interaction. That is, Experiment 1 suggests that with a 5-item list only part of the similarity effect under suppression has an acoustic, as opposed to some other, possibly phonological, origin.

The residual PSE found in primacy could plausibly be ascribed to the action of the phonological store. However, in the remaining two experiments we examine an alternative possibility. It may be the case that the procedural step of using shorter lists produced a different outcome from the study of Jones et al. (2004) not by promoting the use of a phonological store but, rather, by promoting the degree to which perceptual processing could be exploited to recover the order of the to-be-remembered items. To elaborate, with short lists of just 5 auditory items, it is conceivable that automatic order processing may yield useful information about the order of early as well as late items. That is, with a short auditory sequence, the list-initial boundary may help disambiguate the order of the first few dissimilar items just as the end boundary does so for the last few dissimilar items (Jones et al., 2004; present experiment). In Experiment 2, we explore this possibility by introducing a prefix, that is, a redundant item prior to the list (in addition to a suffix). The rationale here is that if the residual PSE found at primacy under suppression (even in the presence of a suffix) is mediated by acoustic coding of the order of items near the list-initial boundary, then occupying the initial as well as the end boundary should now abolish the PSE throughout the list.

Adding a redundant prefix to the to-be-remembered list, even though the prefix is not to be recalled, depresses serial recall performance appreciably (Dallett, 1964, 1965). There is already evidence that this prefix effect is functionally similar in its action to the suffix effect (Jahnke, Nowaczyk, & Wozniak, 1976; Neisser, Hoenig, & Goldstein, 1969) and that it can be used, therefore, as a device to reveal the degree to which acoustic perceptual organization is responsible for the residual PSE found at primacy under suppression in Experiment 1. Several features of the action of a stimulus prefix reveal a resemblance between its action and that of a suffix. In particular, those factors that promote the perceptual partitioning of the prefix from the list diminish its impact, just as is the case for a suffix. So, if the prefix is presented in a male voice and the subsequent list is in a female voice, the damaging effect of the prefix all but disappears (Neisser et al., 1969). Perhaps most telling of all is the fact that increasing the number of prefixes from one to three—the prefixes being three repeated tokens in the same voice as the list—diminishes the prefix effect to a point where prefixed lists are recalled as well as lists without a prefix (Neisser et al., 1969; see Crowder, 1978, for the analogous demonstration in the context of the suffix effect). That is, in terms of our framework of auditory perceptual organization, grouping by similarity partitions the prefixes from the list, restoring the status of the first list item as a boundary to the temporally extended object that is the to-be-remembered list. The result with multiple prefixes (Neisser et al., 1969) also allays concern that the effect of a prefix stems simply from an effect of increasing list length: here, increasing the list length to a length appreciably greater than would be the case with a single prefix, diminishes rather than augments the prefix effect (for a similar finding, see Jahnke et al., 1976).

Experiment 2

Experiment 2 tests the prediction that adding a prefix to the list will ablate the localized residual PSE found under suppression in the presence of a suffix in Experiment 1. We adopt the approach of comparing control lists to lists containing both a prefix and suffix. We predict that now the effect of a prefix will be evident in primacy and the effect of a suffix will be evident in recency. These effects should be additive; that is, the effect of the suffix on recency will be of similar magnitude to that found with a suffix alone in Experiment 1. In other words, the combined action of a prefix and a suffix should abolish the PSE under suppression, pointing to the acoustic, rather than the phonological basis of that similarity effect.

Method

Participants

Twenty-four undergraduate students from the School of Psychology at Cardiff University participated in return for course credits. All were native English speakers reporting 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 1. The only difference was that those lists in Experiment 1 that were followed by a suffix were now also prefixed by the same stimulus as used for the suffix (i.e., the word "*go*" in the same voice as the to-be-remembered list). When present, the onset of the prefix occurred 2 s following the tone that signalled to participants that the next list was imminent. The onset of the first to-be-remembered item followed 750 ms after the offset of the prefix, that is, the prefix was in rhythm with the list. This meant that, regardless of the presence of a prefix, the first to-be-remembered item occurred 3 s after the tone rather than 2 s as was the case in Experiment 1. Including an (optional) 10 min break between blocks, the experiment lasted approximately 70 min.

Results

Fig. 3 shows the general pattern obtained (i.e., without respect to serial position). As predicted by the perceptual organization account, it is clear that under suppression the presence of a prefix and suffix has completely removed the PSE that otherwise survives suppression. Figs. 4A and B show the data across the five serial positions for the 'no suppression' and 'with suppression' conditions, respectively. The presence of a prefix and suffix has served to abolish the PSE throughout the serial position curve.

For the statistical analysis, we again adopt the approach of partitioning the data according to suppression condition. Turning first to the data in Fig. 4A, an ANOVA showed main effects of Similarity, $F(1, 23) = 58.17$, $MSE = 7.78$, $p < .001$, Serial position, $F(4, 92) = 41.60$, $MSE = 2.15$, $p < .001$, and Prefix-and-suffix, $F(1, 23) = 16.89$, $MSE = 2.85$, $p < .001$. There was also a reliable interaction between Similarity and Serial position, $F(4, 92) = 10.99$, $MSE = 1.17$, $p < .001$, reflecting the fact that the PSE tends to be larger toward later serial positions. Furthermore, the three-way interaction between Similarity, Prefix-and-suffix, and Serial position was significant, $F(4, 92) = 5.52$, $MSE = 0.95$, $p < .001$, which may reflect the fact that the tendency for a greater PSE for later serial positions is attenuated in the presence of a prefix-and-suffix. There were no other reliable effects.

Turning to Fig. 4B, an ANOVA showed main effects of Similarity, $F(1, 23) = 17.47$, $MSE = 5.62$, $p < .001$, Serial position, $F(4, 92) = 48.83$, $MSE = 3.38$, $p < .001$, and Prefix-and-suffix, $F(1, 23) = 24.70$, $MSE = 5.23$,

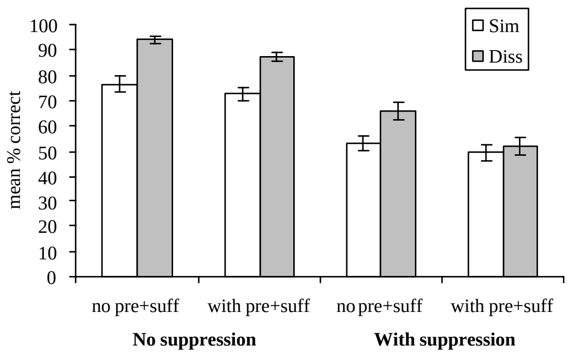

Fig. 3. Mean percentage of phonologically similar (Sim) and dissimilar (Diss) items correctly recalled (without respect to serial position) with and without a prefix-and-suffix (pre + suff) in the 'no suppression' and 'with suppression' conditions in Experiment 2. Error bars represent the standard error of the mean.

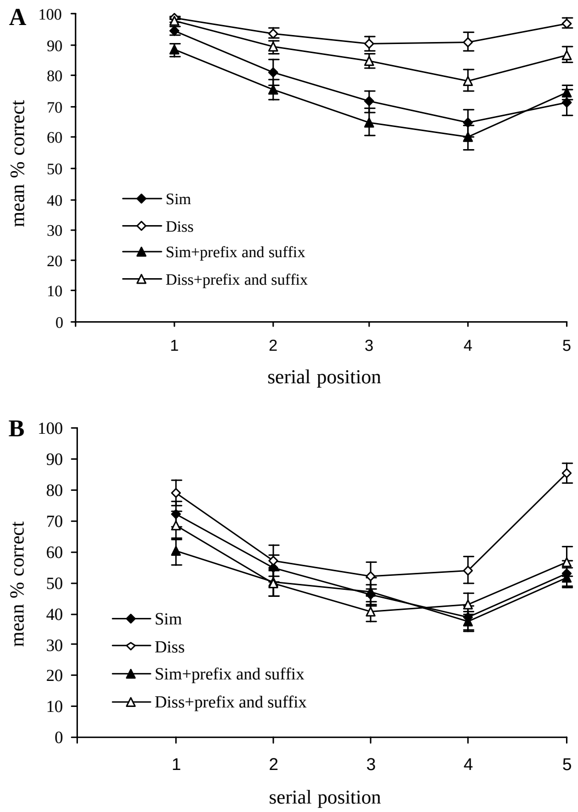

Fig. 4. Mean percentage of phonologically similar (Sim) and dissimilar (Dss) items correctly recalled across serial positions with and without a prefix-and-suffix (pre + suff) in the 'no suppression' (A) and 'with suppression' (B) conditions in Experiment 2. Error bars represent the standard error of the mean.

$p < .001$. All possible interactions were reliable (to at least the $p < .05$ level) including, most critically, the Similarity by Prefix-and-suffix interaction, $F(1, 23) = 5.13$, $MSE = 8.54$, $p < .05$. This interaction was scrutinized further using simple main effects analyses which revealed that the PSE was significant in the 'no prefix-and-suffix' conditions ($p < .001$) but, as predicted by the perceptual organization account, did not approach significance in the 'prefix-and-suffix' conditions ($p = .38$).

Discussion

The results of Experiment 2 are unequivocal: under conditions of articulatory suppression, the presence of a prefix and suffix completely removed the PSE. This result supports the hypothesis, based on the perceptual organization account, that the survival of the PSE at pre-recency in Experiment 1 was due to an opportunistic use of acoustic order encoding for early list items. On the basis of the 'with suppression' data—which provide the best estimate of the contribution made by auditory order processing uncontaminated by rehearsal processes—the prefix and suffix seem to act upon separate portions of the list in similar but discrete ways. The action of the suffix seems to be contained in the recency portion of the serial position curve. Note that the magnitude of the decrease in performance for dissimilar items at recency in the prefix-and-suffix condition in this experiment was almost identical to that found with only a suffix in Experiment 1: averaging performance across the final two serial positions, the suffix alone (Experiment 1) depressed performance for dissimilar items by 19.97% whilst the suffix in conjunction with a prefix (Experiment 2) depressed performance by 19.63%. Thus, the prefix seems not to have added to the effect in recency but is, rather, confined to primacy.

It is far from clear how the phonological storage account could accommodate the fact that the PSE was completely removed under suppression by the presence of a prefix and suffix. However, proponents of the phonological storage account might contend that including both a prefix and a suffix brought the effective "list length" back up to 7 items and thus re-introduced the likelihood that participants would, at least under suppression, abandon the use of the phonological store. Given that this store abandonment argument has been deployed several times as a post hoc means of excluding data that otherwise undermine the phonological store construct (e.g., Baddeley, 2000a, 2000b; Baddeley & Larsen, 2003; Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987; Salamé & Baddeley, 1986), it seems worth addressing it in some detail here.

The store abandonment argument, although a convenient device, is far from compelling both as it might be applied to the present data and in general. First, the number of items that are actually to be recalled is of course not increased by adding a prefix or suffix and indeed, as noted earlier, increasing the number of prefixes (Neisser et al., 1969) or suffixes (e.g., Crowder, 1978) actually enhances rather than depresses serial recall performance. Second, all previous studies of the interaction between similarity, modality, and suppression that have included 7-item lists have shown a robust PSE under suppression with this list length and in these cases none of the items were redundant (i.e., all 7 were to-be-remembered items; Baddeley et al., 1984; Jones et al., 2004; Murray, 1968). Of course, based on the results of Jones et al. (2004) we would expect that this PSE was confined to recency (as noted earlier, no serial position data were reported in Baddeley et al., 1984; we return to discuss the results of Murray, 1968, in General discussion). Nevertheless, the store abandonment argument cannot explain why no PSE was found at recency in the present experiment. A more conservative option for proponents of the phonological storage account would be to concede that the PSE at recency is due to acoustic factors and only the PSE at primacy is due to phonological storage. It could then be argued that whilst the suffix in Experiments 1 and 2 removed the acoustic-based PSE at recency, the addition of the prefix removed the PSE at primacy not because it rendered the use of acoustic coding ineffectual but because it increased the "list length" from 6 to 7 items. However, this more conservative argument is also less than convincing. Note that whether a given list included a prefix-and-suffix was determined pseudo-randomly within the block of trials. It would therefore have to be assumed—somewhat implausibly in our view—that participants adopted a block-wide strategy of selectively and consistently switching to non-phonological coding the moment they heard a prefix.

A third and more general difficulty with the phonological store abandonment argument—regardless of the particular data to which it might be applied—is that it is based ultimately on an unassailable logic. The argument rests on the pre-theoretical assumption that the PSE is indicative of the involvement of phonological storage in short-term verbal recall and hence, by extension, if no PSE is observed then this indicates that the store was not used (e.g., it was abandoned due to too long a list length). This means that the phonological storage construct is irrefutable: the absence of the empirical signature of the store—the PSE—can never be used to question whether such a store exists. The problem is that there is no independent empirical referent for the abandonment of the use of the phonological store and, until such an index is identified, the abandonment argument is a circular one. One counterargument might be that the PSE seems to disappear whenever recall performance drops below some certain level and thus that this drop provides an independent index of phonological store abandonment (see Baddeley, 2000b). This counterargument also fails because it is the increased difficulty of the task—and hence the corresponding drop in performance—that is said to be the *cause* of phonological store abandonment; it cannot at the same time, therefore, also be interpreted as an *effect* of (or empirical referent for) store abandonment. On logical grounds alone, then, it would seem that the onus is on proponents of the phonological store construct to demonstrate empirically—other than by pointing to the disappearance of the PSE—that the various results that have been explained away using the phonological store abandonment argument are indeed due to such abandonment.

We have argued that a prefix and suffix impact upon acoustic rather than phonological processes based on a large body of previous work (Crowder, 1971; Frankish & Turner, 1984; Frick, 1988; Greenberg & Engle, 1983; Greene, 1991; Jahnke et al., 1976; Morton et al., 1971; Neisser et al., 1969; see Nicholls & Jones, 2002a, for a detailed critique of research purporting to show that the suffix effect can be related to post-categorical and not just pre-categorical processes, e.g., Ayres, Jonides, Reitman, Egan, & Howard, 1979; Neath, Surprenant, & Crowder, 1993). Nevertheless, it would seem prudent to demonstrate the acoustic action of the prefix and suffix in the present context. We do this in Experiment 3 by replicating the procedure of Experiment 2 but with one small but potentially critical alteration. Specifically, the word "*go*" at the start and end of the prefix-and-suffix lists was on this occasion presented in a different (female) voice to the (male spoken) to-be-remembered list. Thus, the phonological content of the prefix and suffix was identical to that in Experiment 2; only a pre-categorical, acoustic, characteristic was changed. We predict now that the large acoustic difference between the prefix and suffix on the one hand and the to-be-remembered items on the other will allow the to-be-remembered list to emerge once again as a distinct and independent temporally extended perceptual object (or stream). In other words, the large acoustic difference between a different voice prefix-and-suffix and the list should act in a functionally similar way to the silent periods that usually flank a to-be-remembered auditory list and allow automatically- and acoustically-based order cues to be exploited in the service of serial recall.

Experiment 3

If the abolition of the PSE in the presence of a prefix-and-suffix in Experiment 2 occurred as a result of the fact that the prefix and suffix were acoustically similar to—and hence perceptually grouped with—the to-be-remembered items, decreasing their acoustic similarity to the list should render them far weaker in terms of their power to abolish the PSE under suppression. Thus, in Experiment 3 the survival of the PSE should be evident despite the presence of a (different-voice) prefix-and-suffix. Note that we would still expect to see some reduction in the PSE with a different-voice prefix-and-suffix (compared with no prefix-and-suffix conditions). This expectation is based on the fact that a prefix or suffix in a different voice from a to-be-remembered list typically reduces but does not always abolish their respective power to disrupt serial recall (Greenberg & Engle, 1983; Neisser et al., 1969). From the standpoint of a perceptual organization account, this would be expected on the grounds that a different voice suffix would still bear some acoustic similarity to the to-be-remembered items—e.g., common timing and duration, a similar amplitude envelope, and some similar spectral characteristics—thus precluding its complete perceptual partitioning from the to-be-remembered list. Nevertheless, the key prediction for Experiment 3 is that a reliable PSE will be evident despite the presence of a (different voice) prefix-and-suffix.

Method

Participants

Twenty-four undergraduate students from the School of Psychology at Cardiff University were awarded course credits or received a small honorarium for their participation. All were native English speakers reporting normal hearing and normal or corrected-to-normal vision.

Materials, design, and procedure

These aspects of the methodology were identical to those in Experiment 2 except for the following: the word "*go*" was recorded in a female voice and edited to last 250 ms as for the male spoken "*go*" used in Experiments 1 and 2. Using the "pitch shift" function in the *Sound-Forge* 7 software, the pitch of the token was then shifted up by 2 semi-tones (without altering its duration) to further accentuate the acoustic difference between it and the male spoken to-be-remembered items.

Results

Fig. 5 shows the percentage of correctly recalled items in each of the eight conditions (2 [suppression] by 2 [similarity] by 2 [prefix-and-suffix]) without respect to serial position. It is evident that a PSE was found in all pairs of conditions, most notably, the PSE survived suppression even in the presence of a (different voice) prefix-and-suffix. Figs. 6A and B show the data across the five serial positions for, respectively, the 'no suppression' and 'with suppression' conditions. In line with the perceptual organization account, it is clear from Fig. 6B that the abolition of the PSE under suppression by a prefix-and-suffix found in Experiment 2 was not found here when the prefix and suffix were in a different

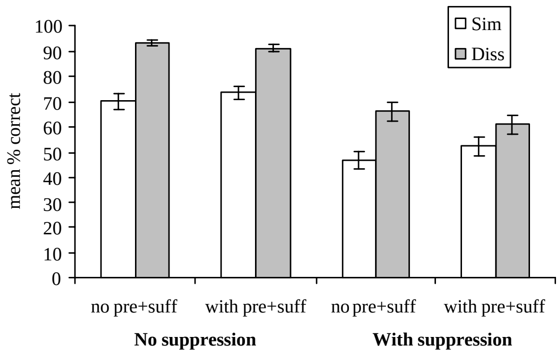

Fig. 5. Mean percentage of phonologically similar (Sim) and dissimilar (Diss) items correctly recalled (without respect to serial position) with and without a (different voice) prefix-and-suffix (pre + suff) in the 'no suppression' and 'with suppression' conditions in Experiment 3. Error bars represent the standard error of the mean.

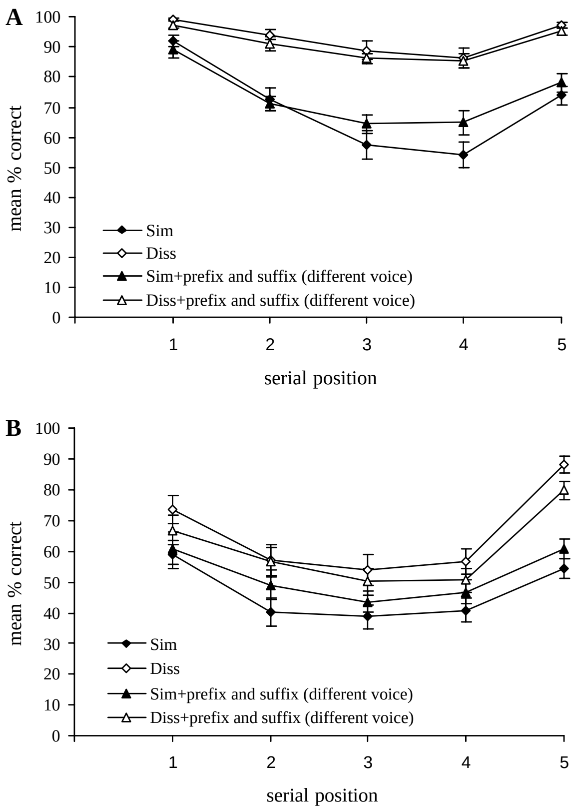

Fig. 6. Mean percentage of phonologically similar (Sim) and dissimilar (Dss) items correctly recalled across serial positions with and without a (different voice) prefix-and-suffix (pre + suff) in the 'no suppression' (A) and 'with suppression' (B) conditions in Experiment 3. Error bars represent the standard error of the mean.

voice from the to-be-remembered list. However, as we also expected, the (different voice) prefix and suffix still seem to have markedly reduced the PSE under suppression.

Turning first to the 'no suppression' data shown in Fig. 6A, an ANOVA showed main effects of Similarity, $F(1, 23) = 141.84$, $MSE = 4.98$, $p < .01$, Serial position, $F(4, 92) = 57.88$, $MSE = 1.88$, $p < .01$, but no main effect of Prefix-and-suffix, $F < 1$. There was a reliable interaction between Similarity and Serial position, $F(4, 92) = 16.13$, $MSE = 1.28$, $p < .01$, again reflecting the fact that the PSE tended to be larger for later portions of the serial position curve. Moreover, the interaction between Prefix-and-suffix and Serial position was significant, $F(4, 92) = 3.73$, $MSE = .85$, $p < .01$, the source of which might be the better recall of similar items in the prefix-and-suffix condition at serial positions 3 and 4. In line with this, the three-way Similarity × Serial position × Prefix-and-suffix also approached significance, $F(4, 92) = 2.32$, $MSE = 1.07$, $p = .063$. There were no other reliable effects.

An ANOVA on the data portrayed in Fig. 6B revealed main effects of Similarity, $F(1, 23) = 75.63$, $MSE = 4.41$, $p < .01$, Serial position, $F(4, 92) = 43.89$, $MSE = 3.66$, $p < .01$, but no main effect of Prefix-and-suffix, $F < 1$. There was again a reliable interaction between Similarity and Serial position, $F(4, 92) = 11.66$, $MSE = 1.47$, $p < .01$, reflecting the fact that the PSE was particularly large for the last serial position (regardless of prefix-and-suffix condition). Finally, the interaction between Similarity and Prefix-and-suffix was significant, $F(1, 23) = 11.11$, $MSE = 4.45$, $p < .01$, reflecting the fact that the survival of the PSE under suppression was far less emphatic in the presence of a (different voice) prefix-and-suffix. Note that this reduction of the PSE seems to have been due in part to a facilitation of performance for similar items in the presence of a prefix-and-suffix, not just an attenuation of performance for dissimilar lists in the presence of a prefix-and-suffix (see following Discussion). No other terms were significant. Simple effects analyses confirmed that although the PSE was diminished by the prefix-and-suffix, the PSE was nevertheless significant both without a prefix-and-suffix and in the presence of a prefix-and-suffix (both $p < .01$).

Discussion

The results of Experiment 3 indicate that the power of a prefix and suffix to abolish the PSE under suppression with auditory lists (Experiment 2) is removed merely by presenting them in a different voice from the to-be-remembered items. This finding is consistent with the contention that the survival of the PSE under suppression at primacy in Experiments 1 and 2 (in the absence of a prefix in the latter case) is based on participants' use of acoustic-based order encoding rather than phonological coding to recall the order of items near the list-initial boundary. When the prefix and suffix are perceptually grouped with the to-be-remembered items (i.e., same voice prefix-and-suffix condition; Experiment 2) acoustic-coding is rendered ineffectual because the order of the items is made ambiguous by the fact that those items are now remote from a boundary. However, when the prefix and suffix are in a different voice from the list, the prefix and suffix are automatically partitioned perceptually from the to-be-remembered list leaving the boundaries of the to-be-remembered list relatively intact.

An intriguing aspect of the data from Experiment 3 is the tendency for the recall of similar items (in both 'no suppression' and 'with suppression' conditions) to have benefited relative to that for dissimilar items from the presence of a (different voice) prefix and suffix. That is, it seems as though the recall of similar items—but not dissimilar items (the recall of which is depressed somewhat due to the different-voice prefix and suffix)—is facilitated when they are flanked by redundant items spoken in a different voice compared to when they are flanked by periods of silence. This is a surprising finding and one that would clearly merit the focus of a separate study. Note, however, that this aspect of the data does not affect the key result of Experiment 3, namely, that the PSE survives suppression in the presence of a prefix and suffix when that prefix and suffix are in a different voice from the to-be-remembered list.

The results of Experiment 3 also serve to further undermine a store abandonment explanation of the disappearance of the PSE under suppression with the same-voice prefix-and-suffix in Experiment 2 (see Discussion of Experiment 2). To re-cap, according to this argument, participants are thought to abandon the phonological store when recall accuracy drops below some certain level. However, in Experiments 2 and 3 the PSE was, respectively, absent and present in the prefix-and-suffix condition despite there being no difference in the general level of recall across the two experiments. Specifically, the level of recall for similar items (collapsed across serial position and across the two prefix-and-suffix conditions) in Experiment 2—in which no PSE was obtained in the presence of a prefix and suffix—was 51.3% correct. However, the level of recall for the same conditions in Experiment 3—in which a PSE was indeed obtained regardless of the presence of a (different voice) prefix and suffix—was, if anything, slightly lower (49.6%). The presence of a PSE in the prefix-and-suffix conditions in Experiment 3 and its absence in these same conditions in Experiment 2 cannot therefore be explained in terms of the hypothesis that participants abandon the phonological store when recall accuracy drops beyond some certain level.

General discussion

We first summarize the pattern of results and begin to spell out their implications for the phonological store construct. In line with the working memory model, Experiment 1 showed that the PSE with 5-item auditory lists under articulatory suppression is evident both in primacy and recency and that a suffix only removes the PSE at recency. However, Experiment 2 indicated that this residual PSE in primacy, like the PSE in recency, appears to be mediated by acoustic not phonological coding. Here, when lists had both a prefix and suffix, the PSE was removed completely. Finally, the results of Experiment 3 served to corroborate the notion that the prefix and suffix disrupted acoustic not phonological coding: changing an acoustic attribute of the prefix and suffix such that they were conveyed in a different voice from the to-be-remembered list, without altering their phonological character, removed their power to abolish the PSE under suppression, although, notably, they still produced a significant reduction of the PSE. The overall pattern of results, then, favours a perceptual organization based account over a phonological storage account of the survival of the PSE under suppression with auditory lists.

Before embarking on an account of the more general implications of the results, we digress briefly to consider a detail associated with the only other instance in which serial position data were reported in the context of the interaction between similarity, modality, and suppression. This is a study by Murray (1968) which was rather brushed aside by Jones et al. (2004) on the grounds that it used probed recall—wherein participants are re-presented with one item from the list and are required to recall the items that followed it—rather than the more standard serial recall task. However, even Murray's (1968) results showed that the survival of the PSE under suppression with auditory presentation is localized at recency: "... recall of the *final items* [emphasis added] of non-articulated auditory lists *is* affected by AC [acoustic confusability][2]...the effects of AC were also marked on early items *but only when rehearsal was permitted*" [emphasis added] (p. 683). Results are highly consistent, therefore, across the three studies for which serial position data are available: Murray (1968), Jones et al. (2004), and the current series.

The present findings illustrate the point that a specifically phonological level of representation is unnecessary for explaining the particular interplay of modality, similarity, and suppression. Additionally, the present results, coupled to those of Jones et al. (2004), suggest that the character of serial recall does not change fundamentally as list length increases (at least in the range of 5–7 items); the change is one of degree rather than kind. Although the PSE under suppression is apparent throughout the serial position curve with shorter lists, this is due to the increasing dominance of auditory effects in primacy, not the engagement of a phonological process. At the same time, the results shed light on the nature of the modality effect and on the contribution made by processes of perceptual organization to auditory short-term recall. As we shall discuss further below, the results point to the conclusion that there is no need to posit a bespoke phonological store for verbal sequences. Instead, these and other results point to a

[2] As noted in Introduction, the term 'phonological similarity effect' has typically been used in favor of 'acoustic confusability effect' on the grounds that visually presented (and hence non-acoustic) as well as auditorily presented items give rise to similar effects. However, as elaborated later in General discussion, our view is that it is the term "phonological similarity" that is a misnomer and that *acoustic* similarity is indeed one of two non-phonological factors that give rise to the illusion of a "phonological" similarity effect.

rather different model of serial short-term memory than one embodied in the working memory framework, one involving two elements: an articulatory (or more generally, gestural) process and a perceptual one.

As noted in Introduction, the general form of the interaction of similarity, modality, and suppression was instrumental in the conception of the phonological loop component of the working memory model (Baddeley, 1986). The fact that the specific form of the interaction turns out to be problematic for the model therefore raises doubts as to the veracity of the phonological loop construct as a whole (including the phonological store). In fact, this is but one thread of evidence that calls the phonological store construct into question. Two major pillars of support for the construct—the PSE and the irrelevant sound effect—can now more accurately and parsimoniously be interpreted by other means. Here, we summarize those arguments relating to the PSE that were covered in greater detail in Jones et al. (2004), and also elaborate some new points in light of the current results (for arguments pertaining to the irrelevant sound effect, see Jones et al., 2004).

''Phonological'' similarity effect?

How does the PSE arise, if not through confusion among tokens within the phonological store? In addressing this question it is useful to distinguish two types of ''phonological'' similarity effect, neither of which are in fact phonological—in the sense of referring to an abstract representation—at all. One is related to the act of rehearsal, the attribution to rehearsal being pointed to emphatically by the loss of the PSE when serial recall is undertaken during articulatory suppression. The other ''phonological'' similarity effect is a product of auditory perceptual organization which contributes in several ways—as observed in the foregoing experiments—to shape performance under suppression with auditory lists. Thus, it can be seen that in both contexts the term ''phonological similarity'' is a misnomer: one is an ''articulatory similarity'' effect and the other an ''acoustic similarity'' effect.

When rehearsal is permitted, the chief way by which the PSE arises is through the very act of rehearsal itself; these are errors akin to Spoonerisms in which elements of items in the to-be-recalled sequence are transposed on the basis of their articulatory similarity. Several authors have made such a suggestion but perhaps the most convincing and complete account is due to Ellis (1980) who showed that the pattern of similarity-based errors in short-term memory was identical to that found when participants were simply required to read lists of phonologically similar items with no memory load (see also Mackay, 1970; Shattuck-Hufnagel & Klatt, 1979). Moreover, the pattern of errors is broadly similar to that found in natural speech production. In addition, the mere fact that when rehearsal is prevented the PSE disappears for both visual and auditory material—so long as the products of auditory perceptual organization are rendered obsolete with the use of suffixes or/and prefixes in the latter case—is compelling evidence of the intimate involvement of speech production processes.

It could be counterargued that phonological similarity errors are far more prevalent in a serial recall setting than in a reading task and thus speech errors cannot entirely account for the PSE (see, e.g., Baddeley & Larsen, 2003). However, the serial recall task is one that —compared to reading or natural speech production—demands a high degree of output planning (usually speech planning in the case in which the to-be-remembered material is verbal) and it is the planning of speech rather than its production per se that is the source of speech errors (e.g., Levelt, 1989). The serial recall setting is one in which, by design, there is a mismatch between the environmental input and extant knowledge (e.g., '*7, 5, 3, 1, 4, 2, 6*' might be presented, but '*1, 2, 3, 4, 5, 6, 7*' would not). According to Neumann (1987, 1989, 1996), it is precisely in the face of such a mismatch—an *underspecification of action parameters* problem in his terms—that output planning is needed. Thus, we identify rehearsal with the planning of a gestural (e.g., articulatory) *sequence* rather than—as is the case within the working memory framework—a vehicle for revivifying individual representations of to-be-remembered items (or their constituent phonemes) whilst some separate competitive queuing process determines the sequence in which those items will be outputted (e.g., Burgess & Hitch, 1999; Henson, 1998; Page & Norris, 1998).

The other ''phonological'' similarity effect arises from the action of acoustic similarity on the perceptual organization of the list. As noted earlier, one way of construing a sequence of auditory stimuli, such as those presented for serial recall, is as a temporally extended auditory object. A range of auditory perceptual factors may modify considerably the capacity to encode order. Not all these phenomena are evident under the normal circumstances of serial recall: performance on the early part of the lists is dominated by rehearsal, whereas at the end of the list passive listening is more the rule, and it is here that the action of auditory factors is typically revealed. When rehearsal is not involved, such as when articulation is suppressed—as was shown in the current series—auditory factors become evident in primacy also, at least with short lists.

One powerful means of disambiguating the order of an auditory sequence is to provide strong boundaries to act as anchors for order retrieval (see Bregman, 1990; Warren, 1982). The silence at the end of a list is one such boundary, and we argue that the recency that typically accompanies auditory presentation is a consequence of its action. Adding a suffix, particularly one sharing physical characteristics with the list, extends the object, the boundary now being signified by the silence following the suffix, not the last list item. Since the suffix is irrelevant to the list, it cannot serve to disambiguate the order of events at the end of the list and the benefit of auditory presentation is lost.

Silence helps to establish the distinctiveness of the position of the last item. However, if the last to-be-recalled item is already relatively acoustically indistinct in relation to its neighbors in the list (e.g, *"b c d g"*), the silence will be ineffectual in enhancing order information and promoting recency. We should expect, therefore, recency to be more pronounced with "phonologically" dissimilar lists than with "phonologically" similar lists, and this appears to be the case (Crowder, 1971; see also Fig. 6A of Jones et al., 2004[3]). The loss of acoustic distinctiveness caused by a suffix is possible only if there is distinctiveness to lose. Under suppression, this strong recency and the advantage for the recall of dissimilar over similar items in recency is particularly pronounced (Jones et al., 2004, Figs. 4B and 6B; present Figs. 2B, 4B, and 6B), indicating that the recall of the last few items in an auditory list is mediated neither by phonological nor articulatory coding (see also Surprenant, Neath, & LeCompte, 1999) but rather by the legacy of perceptual organization processes.

The current experiments have revealed that under articulatory suppression there are effects analogous to the suffix effect at the beginning of the list also. However, in ordinary circumstances of serial recall of auditory lists—in which the list is usually of a greater length than used here (e.g., 6 or more items) and participants are free to rehearse—those aspects of serial recall performance that reflect rehearsal processes will dominate and obscure the high acoustic salience of the first few items (recall of the first item is, in any case, typically at ceiling). In other words, the effect of the initial boundary becomes evident only if rehearsal is prevented—thereby serving simultaneously to force participants to rely more on acoustic coding and to bring performance away from the ceiling—and if the list-initial boundary is displaced by a prefix. This would account for why initial boundary (or primacy) effects are typically not evident whilst end boundary (or recency) effects are a classical signature of auditory serial recall performance (Penney, 1989).

In sum, the interrelation between the action of a suffix or prefix and the PSE is explained by supposing that the distinctiveness that is lent to phonologically dissimilar sequences by the end- or initial-list boundary is sharply diminished if a new boundary is introduced. By definition—given that we equate rehearsal with gestural (e.g., articulatory) planning—the gestural-planning-based similarity effect is only observable when rehearsal is permitted.

Differences from the working memory model

Finally, how distinctive is our approach from that of the working memory model? First, our general level of description differs: instead of stores we use process-based descriptions and constructs such as perceptual organization and output planning. Second, we suppose that there is a distinct perceptual, not phonological, level of representation, to which both speech and non-speech sounds have access. We share with the working memory model the notion of obligatory access but we do not invoke the notion of bespoke stores, rather, we prefer to view this as a process of auditory perceptual organization into streams. Third, we view the representation of verbal rehearsal as being distinctly articulatory, not phonological (as in the "phonological" loop). Fourth, this rehearsal process does not have a reciprocal relation to the perceptual process as in the feedback to the phonological store through the process of revivification. Instead, the perceptual process feeds directly into the output planning process, or rather the output planning process can "pick up" residual information from the perceptual process (though feedback is possible via hearing if there is vocalization of the plan). The process of rehearsal reflects the actuation and revivification of the output plan, not the revivification of some separate representation such as that within the phonological store of the working memory model.

Our perspective is allied more closely to accounts that conceive of short-term verbal memory as being intimately related to language and speech processes (e.g., Gupta & MacWhinney, 1997; MacDonald & Christiansen, 2002; Martin and Saffran, 1997; Martin, Lesch, & Bartha, 1999) than it is to the working memory model. However, several of these language-based accounts only differ from the working memory model insofar as they contend that the short-term memory system responsible for retaining verbal information is not isolable from the language system; that is, in contrast to our perspective, they retain the notion that the cognitive architecture is endowed with special mnemonic properties (e.g., phonological storage; Gupta & MacWhinney, 1997; a phonological buffer prone to decay and interference; Martin et al., 1999; activation and decay; Martin and Saffran, 1997).

Closer to our perspective is that of MacDonald and Christiansen (2002) who eschew temporary storage structures and processes and argue that 'linguistic working memory tasks are simply special kinds of language-

[3] That such a pattern was not evident in the 'no suppression' conditions in the present study (Figs. 2A, 4A, and 6A) is likely due to the fact that performance was near ceiling for dissimilar items at recency (due, in turn, to the use of short lists).

processing tasks' (p. 49; but see Caplan & Waters, 2002). We are also in agreement with MacDonald and Christiansen (2002) in not placing any particular importance on the language system per se in short-term retention. Rather, it *just so happens* that when a short-term 'memory' task involves verbal materials, there exists a particularly rich and multifarious set of skills and habits involved in language-use that lends itself very well to the job of reproducing the order of the stimuli (Macken & Jones, 2003). Thus, the use of language skills (e.g., speech) merely constitutes a restricted example of a general strategy of co-opting motor skills to meet the demands of a short-term 'memory' task (see Jones et al., 2004, for further discussion).

In conclusion, the dominant view of verbal short-term memory supposes that empirical phenomena gleaned from short-term verbal memory tasks reflect the action of a special short-term store dealing with abstract, phonological, representations of verbal events (e.g., Baddeley, 1986, 2000a; Burgess & Hitch, 1999; Page & Norris, 1998). We suggest that such a view constitutes a reification (see, e.g., Macken & Jones, 2003; see also Glenberg, 1997). Very many lines of evidence now converge on the conclusion that such phenomena can more appropriately and parsimoniously be described as being parasitic on a range of what may be regarded as peripheral processes, those associated with the perceptual system (the ones explored experimentally in the current study) and those arising from the operation of the gestural system (revealed by the analysis of speech errors, particularly those arising from "phonological" similarity; see also Murray & Jones, 2002; Woodward et al., 2005). In the case of serial-verbal short-term memory phenomena at least, an abstract phonological store seems entirely superfluous.

Acknowledgments

This research was undertaken with the financial support of the United Kingdom's Economic and Social Research Council (ESRC). Robert Hughes worked on the project whilst an ESRC Post-Doctoral Fellow. Thanks are due to Phillip Morgan for collecting the data and to Alastair Nicholls and Clive Frankish for useful discussions.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.). *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.

Ayres, T. J., Jonides, J., Reitman, J. S., Egan, J. C., & Howard, D. A. (1979). Differing suffix effects for the same physical suffix. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 315–321.

Baddeley, A. D. (1966a). The influence of acoustic and semantic similarity on long term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*, 302–309.

Baddeley, A. D. (1966b). 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. (2000a). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423.

Baddeley, A. D. (2000b). The phonological loop and the irrelevant speech effect: Some comments on Neath. *Psychological Bulletin and Review, 7*, 544–549.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist, 7*, 85–87.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (pp. 47–90). New York: Academic Press.

Baddeley, A. D., & Larsen, J. (2003). The disruption of STM: A response to our commentators. *Quarterly Journal of Experimental Psychology, 56*, 1301–1306.

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

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

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

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Pergamon.

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

Caplan, D., & Waters, G. (2002). Working memory and connectionist models of parsing: A reply to MacDonald and Christiansen (2002). *Psychological Review, 109*, 66–74.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsysScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavioural Research Methods, Instruments, & Computers, 25*, 257–271.

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

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429–460.

Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.

Crowder, R. G. (1971). Waiting for the stimulus suffix: Decay, delay, rhythm, and readout in immediate memory. *Quarterly Journal of Experimental Psychology, 23*, 324–340.

Crowder, R. G. (1978). Mechanisms of auditory backward masking in the stimulus suffix effect. *Psychological Review, 85*, 502–524.

Crowder, R. G. (1993). Auditory memory. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 113–145). Oxford: Oxford University Press.

Crowder, R. G., & Morton, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, 365–373.

Dallett, K. M. (1964). Effects of a redundant prefix on immediate recall. *Journal of Experimental Psychology, 67*, 296–298.

Dallett, K. M. (1965). 'Primary memory': The effects of redundancy upon digit repetition. *Psychonomic Science, 3*, 237–238.

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.

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.

Frankish, C. F., & Turner, J. (1984). Delayed suffix effect at very short delays. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 767–777.

Frick, R. W. (1988). The role of memory in attenuations of the suffix effect. *Memory & Cognition, 16*, 15–22.

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

Greenberg, S. N., & Engle, R. W. (1983). Voice changes in the stimulus suffix effect: Are the effects structural or strategic? *Memory & Cognition, 11*, 551–556.

Greene, R. L. (1991). Serial recall of two-voice lists: Implications for theories of auditory recency and suffix effects. *Memory & Cognition, 19*, 72–78.

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

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression, and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin & Review, 10*, 435–444.

Hanley, J. R., & Broadbent, C. (1987). The effects of unattended speech on serial recall following auditory presentation. *British Journal of Psychology, 78*, 287–297.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13*, 644–655.

Henson, R. N. A. (1998). Short term memory for serial order: The start-end model. *Cognitive Psychology, 30*, 73–137.

Houghton, R. J., Macken, W. J., & Jones, D. M. (2003). Attentional modulation of the visual motion aftereffect has a central cognitive locus: Evidence of interference by the post-categorical on the pre-categorical. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 731–740.

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.

Jahnke, J. C., Nowaczyk, R. H., & Wozniak, W. (1976). Stimulus redundancy and immediate recall. *Memory & Cognition, 4*, 357–360.

Jones, D. M., & Macken, W. J. (1995a). Organizational factors in the effect of irrelevant speech: The role of spatial location and timing. *Memory & Cognition, 21*, 318–328.

Jones, D. M., Macken, W. J., & Harries, C. (1997). Disruption of short-term recognition memory for tones: Streaming or interference? *Quarterly Journal of Experimental Psychology. A, 50*, 337–357.

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., Saint-Aubin, J., & Tremblay, S. (1999). Modulation of the irrelevant sound effect by organizational factors: Further evidence from streaming by location. *Quarterly Journal of Experimental Psychology. A, 52*, 545–554.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Mass: MIT Press.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109*, 35–54.

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

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

Martin, N., & Saffran, E. M. (1999). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology, 14*, 641–682.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in short-term memory and word processing. *Journal of Memory and Language, 41*, 3–29.

Morton, J., Crowder, R. G., & Prussin, H. A. (1971). Experiments with the stimulus suffix effect. *Journal of Experimental Psychology, 91*, 169–190.

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.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review, 10*, 435–444.

Neath, I., Surprenant, A. M., & Crowder, R. G. (1993). The context dependent stimulus suffix effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 698–703.

Neisser, U., Hoenig, Y. J., & Goldstein, E. (1969). Perceptual organization in the prefix effect. *Journal of Verbal Learning and Verbal Behavior, 8*, 424–429.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 361–393). Hillsdale, NJ: Erlbaum.

Neumann, O. (1989). On the origins and status of the concept of automatic processing. *Zeitschrift für Psychologie, 197*, 411–428.

Neumann, O. (1996). Theories of attention. In O. Neumann & A. F. Sanders (Eds.). *Handbook of perception and action* (Vol. 3, pp. 389–446). London: Academic Press.

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

Nicholls, A. P., & Jones, D. M. (2002b). Perceptual organization in short-term memory: The sandwich effect reassessed. *Memory & Cognition, 30*, 81–88.

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

Pashler, H. E. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.

Penney, C. G. (1989). Modality effects and the structure of short-term memory. *Memory & Cognition, 17*, 398–422.

Peterson, L. R., & Johnson, S. T. (1971). Some effects of minimizing articulation on short-term memory. *Journal of Verbal Learning and Verbal Behavior, 10*, 346–354.

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

Porter, R. J. (1987). What is the relation between speech production and speech perception? In A. Allport, D. MacKay, W. Pring, & E. Scheerer (Eds.), *Language perception and production: Relationship between listening, speaking, reading, and writing* (pp. 85–106). London, England: Academic Press.

Salamé, P., & Baddeley, A. D. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24*, 263–265.

Shattuck-Hufnagel, S., & Klatt, D. (1979). The limited use of distinctive features and markedness in speech production: Evidence from speech error data. *Journal of Verbal Learning and Verbal Behavior, 18*, 41–55.

Surprenant, A. M., Neath, I., & LeCompte, D. C. (1999). Irrelevant speech, phonological similarity, and presentation modality. *Memory, 7*, 405–420.

Warren, R. J. (1982). *Auditory perception: A new synthesis*. Oxford: Pergamon.

Warren, R. M., & Obuzek, C. J. (1972). Identification of temporal order within auditory sequences. *Perception & Psychophysics, 12*, 86–90.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review, 8*, 44–57.

Woodward, A., Macken, W. J., & Jones, D. M. (2005). Linguistic familiarity in short-term memory: (Co-) articulatory fluency or redintegration? Manuscript in preparation.