Journal of Memory and Language 90 (2016) 126–146

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

# Perceptual-motor determinants of auditory-verbal serial short-term memory ☆

Robert W. Hughes [a,*], Cindy Chamberland [b], Sébastien Tremblay [b], Dylan M. Jones [c,1]

[a] Department of Psychology, Royal Holloway, University of London, Egham, Surrey, UK
[b] École de psychologie, Université Laval, Québec, Canada
[c] School of Psychology, Cardiff University, Cardiff, UK

ARTICLE INFO

*Article history:*
Received 25 July 2015
revision received 20 April 2016
Available online 8 May 2016

*Keywords:*
Perceptual organization
Motor-planning
Short-term memory
Serial recall
Perceptual variability

ABSTRACT

The role of the compatibility between obligatory perceptual organization and the active assembly of a motor-plan in auditory-verbal serial recall was examined. The classic finding that serial recall is poorer with ear-alternating items was shown to be related to spatial-source localization, thereby confirming a basic tenet of the perceptual-motor account and disconfirming an early account characterizing the two ears as separate input-channels (Experiment 1). Promoting the streaming-by-location of ear-alternating items—and therefore the incompatibility between perceived and actual order—augmented the ear-alternation effect (Experiment 2) whereas demoting streaming-by-location by reducing the regularity of the alternation attenuated it (Experiment 3). Finally, increasing the perceptual variability of an ear-alternating list while demoting the likelihood of streaming-by-location—by adding uncorrelated voice changes—also reduced the ear-alternation effect as did articulatory suppression for that part of the list (pre-recency) associated with motor-planning (Experiment 4). The results are incompatible with theories in which perceptual variability impairs serial recall due to a deficit in encoding items into a limited-capacity short-term memory space and instead point to a central role for perceptual and motor processes in serial short-term memory performance.


## Introduction

The ability to retain and reproduce a sequence of stimuli over the short-term has long been recognized as a fundamental aspect of cognition, playing a critical role in many higher-level functions including problem-solving, reasoning, speech processing, and language learning (e.g., Baddeley, 1986, 2007; Hurlstone, Hitch, & Baddeley, 2014; Lashley, 1951; Rosenbaum, Cohen, Jax, Weiss, & van der Wel, 2007). Classically, the study of verbal short-term memory has been wedded conceptually to the formalist, psycholinguistic, tradition (cf. Chomsky & Halle, 1968) in which the key unit of analysis is modality-independent, phonological, representation (e.g., Baddeley, 2007, 2012; Baddeley & Hitch, 1974; Broadbent, 1958,

☆ Author note: We would like to thank Al Bregman for valuable comments on an early version of the paper. Thanks are due also to Mark Hurlstone for programming Experiments 1 and 3 and Marie-Josée Côte, Laurence Dumont, Ben Jones, and Marie-Kim Côté for data collection. The research was supported in part by an *Economic and Social Research Council* grant awarded to Dylan M. Jones, Bill Macken, & Robert W. Hughes. The results of Experiments 1 and 2 were presented at the Meeting of the *Experimental Psychology Society*, University College London, January, 2014, and at the *International Conference on Working Memory*, Cambridge, July, 2014.
* Corresponding author.
*E-mail address:* Rob.Hughes@rhul.ac.uk (R.W. Hughes).
[1] Also at the University of Western Australia.

http://dx.doi.org/10.1016/j.jml.2016.04.006
0749-596X/

1984; Pashler, 1998). In this view, the capacity to recall a series of verbal items is understood in terms of the operation of a dedicated but highly fragile system or representational space in which central (e.g., phonological) representations of such items outlast their physical presence but are subject to inexorable decay or/and interference from other structurally similar items (Baddeley, 2007). Accordingly, research set within this centralist view has tended to focus on structural factors such as the duration of the short-term representation (e.g., Baddeley, Thomson, & Buchanan, 1975; Barrouillet & Camos, 2014), the particular mechanism by which it might be corrupted by other events occupying the same memorial space (e.g., Neath, 2000; Oberauer, 2002), or the overall capacity of that space (e.g., Cowan, 2001, 2015).

We argue here that an emphasis on the structural properties assumed to underpin short-term memory performance has obscured a key role for general-purpose perception and motor functions that have often been seen as peripheral; merely providing the input to, and output from, central short-term memory mechanisms. Indeed, while some centralist accounts now make strong links between short-term memory mechanisms and perception and action processes (e.g., Page, Madge, Cumming, & Norris, 2007), there is a burgeoning body of work suggesting that a consideration of perceptual organization, motor-planning, and the mapping between them may go a long way to accounting for short-term memory phenomena without having to invoke dedicated storage mechanisms (e.g., Guérard & Tremblay, 2011; Hughes & Jones, 2005; Hughes, Marsh, & Jones, 2009, 2011; Jones, Hughes, & Macken, 2006, 2007; Jones, Macken, & Nicholls, 2004; Macken, Taylor, & Jones, 2014, 2015; Maidment & Macken, 2012; for related views, see Acheson & MacDonald, 2009; MacDonald, 2016; Reisberg, Rappaport, & O'Shaughnessy, 1984; Wilson & Fox, 2007). The aim of the present research is to examine the way in which the passive perceptual organization of the auditory scene into coherent temporally-extended perceptual objects or streams (cf. Bregman, 1990) is a key determinant of the short-term reproduction of a spoken sequence. More specifically, we study the importance of the compatibility between the obligatory, non-volitional, organization of spoken items according to their perceived spatial-source and the active assembly of those items in their true temporal order in (subvocal) articulatory form. As a theoretical counterpoint, we contrast the predictions of our perceptual-motor account with ones allied to the centralist view in which factors such as variation in spatial-source are seen as compromising the initial encoding of items into short-term memory rather than ones that affect processes that are integral to short-term memory performance.

### *Serial short-term memory: a perceptual-motor approach*

Present understanding of serial short-term memory is based primarily on performance in a verbal serial recall task in which, typically, around 5–8 verbal items (e.g., digits, letters, words) are presented one at a time and which must be reproduced in strict serial order following the last item (Baddeley, 1966; Conrad, 1964). A key observation at the heart of the perceptual-motor account is that the post-categorical identities of the items in a serial recall list, by design, exhibit very low transitional probabilities (i.e., the predictability of an item given the preceding event(s) is very low or zero; e.g., Miller & Chomsky, 1963): That is, the Experimenter typically strips the list of supra-item features—syntax, grammar, and semantic relations—that in a normal linguistic sequence constrain the serial order of its constituent elements (see, e.g., Jefferies, Lambon Ralph, & Baddeley, 2004; Macken & Jones, 2003). At least in relatively 'pure' serial recall tasks, in which the burden falls entirely or primarily on reproducing item order rather than individual item identity (cf. Baddeley, 2012), it is this characteristic (alone) that makes serial recall challenging. Accordingly, performance is superior when there is a good match between the list and long-term sequential knowledge (i.e., when the list-items exhibit relatively high transitional probabilities) such as with a list containing high-frequency letter transitions (Miller & Selfridge, 1950), a list of words that make up a grammatically legitimate sentence (Jefferies et al., 2004), or a list that contains subsequences already unitized in long-term memory due to repeated exposure to, for instance, telephone or personal identification numbers (Jones & Macken, 2015).

Given the lack of serial order constraints within the presented list, we argue that the motor-sequence planning system (vocal-articulatory in the case of verbal serial recall) is co-opted to impose such constraints. The skill of speech (or more accurately, speech-planning) provides a particularly effective medium for this purpose on account of its inherent sequentiality, continuity, and prosodic and co-articulatory nature. For example, the very act of (covertly) co-articulating the items—whereby the exact manner in which the end of one speech element is articulated depends on the next speech element (Sternberg, Wright, Knoll, & Monsell, 1980)—generates new sequential information (and hence constraints) not present in the list itself (e.g., Woodward, Macken, & Jones, 2008). Importantly, then, in the perceptual-motor account, the function of articulatory rehearsal (or vocal-motor planning) is not the refreshing of decaying representations within a distinct verbal (phonological) store (or the conversion of visually-presented items into phonological form; e.g., Baddeley, 2007) but rather the motoric sequential binding of otherwise post-categorically unbound items.

A second key facet of the perceptual-motor approach is a consideration of obligatory perceptual processes that organize sensory input into coherent perceptual objects (e.g., Wertheimer, 1923/1938). Of particular relevance are the especially powerful processes of auditory scene analysis that generate objects or 'streams'; representations that, unlike representations of visual stimuli, are inherently sequential given the temporally-unfolding nature of sound (e.g., Bregman, 1990). To elaborate, auditory sequential streaming involves the computation of whether or not successive sounds are emanating from the same environmental event and whether they should therefore be assigned to the same stream (stream integration) or to different streams (stream segregation), respectively. This is widely assumed to be an obligatory, non-volitional, process involving the preattentive processing of a variety of cues