Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 56 (2007) 49–64
Journal of Memory and Language
www.elsevier.com/locate/jml

Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus ☆

Mike P.A. Page [a,*], Alison Madge [a], Nick Cumming [a], Dennis G. Norris [b]

[a] University of Hertfordshire, Department of Psychology, College Lane, Hatfield AL10 9AB, UK
[b] MRC Cognition and Brain Sciences, 15, Chaucer Road, Cambridge, UK

Received 27 July 2005; revision received 14 September 2006
Available online 2 November 2006

Abstract

In three experiments, we tested the hypothesis that those errors in immediate serial recall (ISR) that are attributable to phonological confusability share a locus with segmental errors in normal speech production. In the first two experiments, speech errors were elicited in the repeated paced reading of six-letter lists. The errors mirrored the phonological confusions seen in ISR. In a third experiment, participants performed ISR for four-word lists. Some of the lists were designed to encourage the exchange of onset consonants between adjacent words. ISR was shown to be sensitive to this manipulation, further supporting the common-locus hypothesis. The results are discussed in the context of theories of serial recall and of speech production, and are further related to neuropsychological data.

Keywords: Immediate serial recall; Speech production; Speech errors; Phonological similarity effect; Working memory

In this article, we present evidence that the errors underlying the phonological similarity effect (PSE) in immediate serial recall (ISR) are similar in character to errors seen in spontaneous speech production. We therefore explore the hypothesis that the two types of error result from the operation of a common mechanism. The PSE is characterized by increased errors in the immediate serial recall of lists containing words that sound similar. This effect has been shown in lists of rhyming letters (e.g., "BGTCVP" Baddeley, 1968; Conrad, 1964; Conrad & Hull, 1964; Wickelgren, 1965), and with lists of rhyming words and/or words that share a vowel (e.g., Baddeley, 1966). More recent work (Fallon, Groves, & Tehan, 1999; Nimmo & Roodenrys, 2004) has confirmed that order-memory for lists of CVC words suffers whether items share either rime (vowel and coda), or onset and coda, or onset and vowel. These studies also found that *item* recall can actually improve when items share a rime, probably because the rime is a salient cue that can assist in determining which items were present (though not the order in which they occurred). This enhanced item-memory can mask the PSE when a conventional correct-in-position scoring method is used, though not when order errors are conditionalized on a free-recall measure of the item in question. Nimmo and Roodenrys noted this sensitivity to the rime unit, and more generally the increased influence that vowel similarity exhibited in their data relative to similarity based on shared consonants. They related these factors to models, such as that of Hartley and

☆ We thank Dr. Mike Greenhough, Jane Hall, and Jon Williams for their assistance in running the experiments reported here.
* Corresponding author. Fax: +44 1707 285073.
E-mail address: m.2.page@herts.ac.uk (M.P.A. Page).

0749-596X/$ - see front matter
doi:10.1016/j.jml.2006.09.002

Houghton (1996), that seek to place linguistic constraints on the representations used in both short-term memory and in speech production. Nimmo and Roodenrys concluded that there was "an urgent need for STM researchers to integrate linguistic research, and models based on this research, into STM models". This paper is part of an attempt to do just that.

Other authors have drawn attention to the similarity between verbal STM and speech production. Most notably Ellis (1980), influenced by the earlier contributions of Morton (1964, 1968, 1970), explored the proposal that immediate recall of verbal materials was carried out using a "response buffer". The primary function of this buffer was taken to be the storage of a speech programme during the period between speech planning and overt articulation. Ellis proposed what he called the "error equivalence hypothesis", namely, that if a common response buffer was involved in both speech production and short-term memory for serial order, then similar types of phonemic error would be expected in both tasks. In three experiments, involving recall of lists of CV and VC syllables, he corroborated this hypothesis by showing: that serial recall errors most often involved consonant swaps (more so than either vowel swaps or whole-syllable swaps); that consonant errors respected a feature-similarity effect, such that consonants tended to exchange when they were more featurally similar; that phonemic errors tend to preserve within-syllable position; and, finally, that consonant swaps were more numerous between syllables that shared a vowel, a pattern he dubbed the "contextual similarity effect". All of these serial recall effects had previously been observed in relation to speech production errors (e.g., MacKay, 1970; Nooteboom, 1967).

Although it is not something to which Ellis (1980) drew direct attention, it is the contextual similarity effect that can be applied most directly to the PSE. In a typical phonologically confusable stimulus-list of rhyming single-syllable items, the context similarity effect might be expected to apply with force: there are many onset consonants that share the same context, and this context is not just similar but is identical in both vowel and coda. Moreover, in rhyming lists, any onset-consonant exchange will result in the same items as in the stimulus list, just placed in a different order. Such exchanges are difficult for the speaker to detect because no unintended item is thereby introduced into their recall. The key observation is therefore that, although the extra order errors seen in recall of a list of rhyming items have traditionally been seen as exchanges of complete items, they might perhaps better be thought of as onset exchanges promoted by contextual similarity.

Explaining the PSE in terms of speech production mechanisms has a clear advantage from the perspective of computational models of verbal ISR. In the last decade or so, we (Henson, Norris, Page, & Baddeley, 1996; Norris, Page, & Baddeley, 1994; Page & Norris, 1998a, 1998b) and others (Burgess & Hitch, 1992, 1999; Henson, 1998) have used data from ISR tasks to help develop computational simulations of short-term memory. One of the most important constraints on these models was provided by data from recall of lists of alternating confusability. Such lists, including those in which rhyming items are placed at alternating list-positions (e.g., "BRPXDQ"), are interesting because it has been shown that while the rhyming items are subject to additional recall errors (usually mutual exchanges), the interleaved nonrhyming items are recalled as well as they would be in a list comprised entirely of nonrhyming items. Indeed, Farrell and Lewandowsky (2003), have recently claimed that nonconfusable items are recalled *better* in mixed lists than in pure nonconfusable lists. For the alternating lists, this pattern of errors results in a serial position curve that has a characteristic saw-tooth shape, with error-peaks located at the stimulus-list positions occupied by confusable items. These data are difficult to explain in terms of "chaining" models of ISR, in which each list-item is associatively chained to its predecessor(s) (Henson et al., 1996). Moreover, early position-item association models (e.g., Burgess & Hitch, 1992) were unable to simulate such a pattern. The key factor that enabled the primacy model (Henson et al., 1996; Norris et al., 1994; Page & Norris, 1998a, 1998b), and other later models (Burgess & Hitch, 1999; Henson, 1998) to simulate these data accurately was the incorporation of a two-stage, or two-phase, recall process. The precise details differ between models (see Page & Henson, 2001, for a review), but they all involve an initial stage/phase in which an item is selected on the basis of order information, followed by a second stage/phase, in which the selected item can be replaced at output by one with which it is phonologically confusable. Although the use of a two-stage mechanism allowed the models to simulate the PSE data, there appeared to be no independent motivation for using two stages. The second stage of the primacy model and related models does nothing other than introduce additional errors that would not occur in its absence. While the data seemed to require a second stage, its presence was thus something of an embarrassment. The appeal to parallels between ISR and speech production provides a way out of this somewhat uncomfortable situation.

In Page and Norris (1998a), and more explicitly in a companion chapter (Page & Norris, 1998b), we pointed out that most modern models of speech production are also inherently two-stage in nature. The requirement for two stages in models of speech production follows from the fact that speech is more than just a concatenation of discrete words. For a multiword utterance, once the phonological representation of the words is read from the lexicon, further processes must operate on those representations to produce speech that is fluent. A com-