Running head: PHONOLOGICAL SIMILARITY PI AND STM

Is the phonological similarity effect in working memory due to proactive interference?

Alan Baddeley

Graham Hitch

Philip Quinlan

Department of Psychology, University of York

Address Correspondence to:

Alan Baddeley

Department of Psychology

University of York

Heslington

York YO10 5DD

United Kingdom

Phone: +44 (0)1904 322882

Fax: +44 (0)1904 433181

Email: ab50@york.ac.uk

Abstract

Immediate serial recall of verbal material is highly sensitive to impairment due to phonological similarity. While this has traditionally been interpreted as a within-sequence similarity effect, Engle (2007) proposed an interpretation based on interference from prior sequences, a phenomenon analogous to that found in the Peterson short-term memory task. We use the method of serial reconstruction to test this in an experiment contrasting the standard paradigm in which successive sequences are drawn from the same set of phonologically similar or dissimilar words and one in which the vowel sound on which similarity is based is switched from trial to trial, a manipulation analogous to that producing release from PI in the Peterson task. A substantial similarity effect occurs under both conditions although there is a small advantage from switching across similar sequences. There is however, no evidence for the suggestion that the similarity effect will be absent from the very first sequence tested. Our results support the within-sequence similarity rather than a between-list PI interpretation. Reasons for the contrast with the classic Peterson short-term forgetting task are briefly discussed.

*Key words:* Working memory, short-term memory, serial reconstruction, phonological similarity, proactive interference

One of the most robust and substantial phenomena in the area of short-term and working memory[[1]](#footnote-1) is the deleterious effect of within-sequence phonological similarity on immediate serial recall. Conrad and Hull (1964) showed substantially poorer short-term memory (STM) for acoustically similar letters (b d c p v) than dissimilar (f w k r y). Baddeley (1966a) showed that this effect also applied to word sequences, with sequences such as Man, Cat, Map, Mat, Can much more error prone than dissimilar sequences, while the effects of semantic similarity (eg Huge, Large, Big, Long Tall) were minimal. In contrast , the long-term learning of longer sequences showed exactly the opposite pattern (Baddeley, 1966b). Historically, this effect formed an important part of the evidence for the distinction between short-term and long-term memory stores in the modal model of memory (Atkinson & Shiffrin, 1968). It was also regarded as central to the phonological loop component of the subsequent multicomponent model of working memory , a subsystem that is assumed to be capable of encoding and rehearsing verbal and acoustic material and to play an important role in facilitating long-term phonological learning and hence in the acquisition of native and second language vocabulary (Baddeley, Gathercole & Papagno,, 1998; Baddeley & Hitch, 1974; Baddeley, 2012).

Wickelgren (1965) provided evidence that the effect occurs at the level of retention of serial order, rather than through differential availability of items. Conrad (1965) accounted for this by noting that errors in the recall of individual items tended to be phonologically similar and drawn from the small set of recently presented items. For a sequence of phonologically similar items, a series of such item substitution errors would result in poorer ordered recall. This has continued to be the dominant view with a range of later models of immediate serial recall implementing mechanisms that result in phonologically based confusions among the most recently presented items (Botvinick & Plaut, 2006; Burgess & Hitch, 1992; 1999; Lewandowsky & Farrell, 2008; Page & Norris, 1998).

However, despite this general agreement, an alternative explanation has been proposed by Engle (2007) who suggests that the effect may stem from interference from prior sequences, rather than from within-sequence similarity per se. Engle’s proposal is consistent with the importance placed on inhibition within his working memory model (Engle & Kane, 2004; Shipstead & Engle, 2013). It also links to earlier literature on the Peterson and Peterson (1959) short-term forgetting paradigm whereby a consonant trigram showed dramatic forgetting over an 18s period provided ongoing rehearsal was prevented, an effect that the Petersons attributed to the gradual fading of a short-term memory trace. Subsequently however, Keppel and Underwood (1962) presented strong evidence against this view, demonstrating that forgetting did not occur on the very first trigram, but built up over the first four or five items, interpreting the effect in terms of proactive interference (PI) from earlier trigrams. Further evidence for this interpretation came from the phenomenon known as release from PI whereby changing the category of the material to be retained led to excellent performance on the first item following the change in category, followed by a gradual decline in performance over the subsequent three or four items. This was shown to occur when letters were followed by digits (Wickens, Born & Allen, 1963), while a study by Loess (1968) switched between blocks of word triplets selected from a range of semantic categories, finding that on each occasion, the introduction of a new category enhanced performance on the item following the switch, after which performance gradually declined.

While the release from PI phenomenon is robust, its interpretation in terms of classic interference theory has been questioned. Loess and Waugh (1967) found for example that leaving an unfilled gap between successive items removed the PI effect while an elegant study by Turvey, Brick and Osborne (1970) suggested that some form of temporal discrimination offered a better interpretation of the release effect, a view that is reflected in later proposals ranging from Baddeley (1986) to the much more extensively developed SIMPLE model of Brown, Neath and Chater (2007). The importance of between-list effects was also emphasised by Estes (1991) in a paper concerned with immediate verbal serial recall that focuses on intrusion errors which he shows frequently come from earlier lists, emphasising the importance of intertrial effects proposing what he terms, the trial unit model, in which, “there are no built in boundaries between representations of the events of successive trials” (Estes, 1991 p168). However, as Conrad (1967) points out, while intrusion errors are often interpreted as indicating the source of interference, they may simply be filling a gap due to trace decay with the most available alternative response. Hence, while intrusion errors may be informative, their interpretation is open to question.

At the core of Engle’s (2007) reconsideration of the phonological similarity effect is his proposal that the principal purpose of working memory is to allow us to function in the face of proactive interference. Referring to the Keppel and Underwood (1962) study, he concludes that effects such as that of phonological similarity are “probably found only under conditions of interference”, concluding “I predict that effects such as phonemic similarity, articulatory suppression and other effects of short-term memory studies would also not be found over the first trials in an experiment”.

Somewhat surprisingly, Engle’s interpretation does not appear to have been tested. We describe a simple test of this ingenious re-interpretation of the influence of phonological similarity on STM. Our design involves the immediate serial recall of sequences of six words. These were selected on each trial from different combinations drawn from sets of six phonologically similar CVC words comprising one set from each of six different vowel sounds (see Table 1). The standard condition within the literature involves the same set of phonologically similar or dissimilar items being used repeatedly, hence maximising the potential for between-sequence PI. We compare this with a condition in which PI is minimised by using a different set of similar or dissimilar items for each of six sequences within a test block. According to Engle’s interpretation, performance should be greatly enhanced in this latter “release from PI” condition. Furthermore, across all four conditions, performance on the initial trial should be equivalent, given minimal between-list PI at this point. On subsequent trials, performance in the standard phonological similarity condition should decline, reflecting the build-up of PI when the same similar items are used repeatedly. In contrast, performance over subsequent trials in the “release from PI” condition should remain high and equivalent to performance with dissimilar items, given minimal between-list PI in these conditions.

Method

Table 1 about here

Design

 This involved four conditions as follows in a within-participants design: the term “repeat” indicates that the same six words were used repeatedly for a given participant whereas the term “change” indicates the use of all 36 words.

1. Repeat Similar: This corresponds to the standard paradigm whereby a limited set of similar items is used repeatedly, for example the first sequence might be man, cat, bat, ban, pad, can followed on the next trial by ban, man, bat, pan, cat, pad; always the same items but in a different order. However, although each participant consistently used the same vowel set, different participants were tested using sets based on different vowel sounds (i.e. A to F in Table 1).
2. Change Similar: In this condition participants were tested on items from a different similar set on each of six successive trials (e.g. pan, cat, bat, ban, pad, man followed by hen, bed, jet, pen, fed, pet, etc.).

(3) Repeat Dissimilar: In this case, one item was selected from each of the vowel categories in sets A to F, with the same items being used for each list (e.g. man, pen, rim, cod, bud, peel followed by bud, peel, rim, pen, cod, man, etc.). Different participants were tested with a different selection of items.

Finally,

(4) Change Dissimilar: Here, a different dissimilar set was used on each of six successive trials (e.g. man, pen, rim, cod, bud, peel followed by cat, bin, bug, pet, feel, cot, etc.).

A total of 12 lists were constructed for each condition. Thus in the Change Similar and Change Dissimilar conditions each of the six sets of words was used twice in separate halves of the trial sequence. Subject to the constraints defining each condition, the selection of items and their orders of presentation were freshly randomised for each participant. This study forms part of a programme approved by the University of York Psychology Ethics Committee.

Participants

Participants were twenty six students from the University of York who took part for course credit or a small monetary reward.

Procedure

We used a serial reconstruction method which, by making the relevant items available, emphasises the role of serial order which as shown by Wickelgren (1965) is central to the phonological similarity effect. Items were presented visually on a VDU screen at a rate of one per second with a 250msc inter item interval. The final item was followed by a query which was followed after 1s by a response screen showing the six presented items in a column on the left hand side in a random order. The right hand side displayed a column of empty boxes. Participants had to reconstruct the serial order of the sequence by clicking on each word and moving it to the appropriate serial position marker on the right of the screen. They were encouraged to guess if necessary, and were given the option of entering a blank at any position if they preferred. Retrieval was unpaced.

Each participant completed a block of 12 trials for each condition in a counterbalanced order. There were no practice trials so as to ensure maximum control over PI within the experiment.

Results

The mean numbers of correctly recalled items are shown in Figure 1.

Figure 1 about here

A 2 x 2 repeated measures ANOVA on number correct indicated a substantial effect of similarity, *F*(1, 25) = 55.9, *p*< .01, $η\_{p}^{2}$ = .69, and a weaker effect of repetition of the same items on each trial, *F*(1,25) = 4.83, *p*< .05, $η\_{p}^{2}$ = .16. However, the interaction suggested by inspection of Figure 1 was not significant, *F*(1,25) = 1.74, *p* = .20, $η\_{p}^{2} $= .07.

Planned comparisons were used to test the pre-specified hypotheses. The clearest test of Engle’s hypothesis comes from comparing the standard condition in which the same vowels are used throughout with the classic release from PI condition in which successive lists use different vowels. This does produce a significant difference with repeated vowels resulting in a mean of 2.83 (*SD* = .40) words correct compared to 3.16 (*SD* = .94), *t*(25) = 2.42 *p* = .02 , Cohen’s *d* = 0.36, a small to medium effect. Although indicating a degree of release from PI, however, this does not remove the substantial effect of phonological similarity as reflected in the comparison between the “release” condition and conditions 3 (*M* = 4.21, *SD* = 1.13) and 4 (*M* = 4.27, *SD* = 1.00) which repeated and changed dissimilar vowel sounds, *t*(25) = 5.37, *p*<.001, *d* = 1.01 and *t*(25) = 7.61, *p*< .001, *d* = 1.14 respectively). The absence of a difference between condition (3) in which the same 6 items occur on every trial and condition (4) in which every trial involves different items is also inconsistent with a major contribution from PI (*t* (25) = .43, *p* = .67, *d* = .06).

Trial by trial data for the four conditions are shown in Figure 2. There is no suggestion of the predicted enhancement in performance on trial 1, nor is there evidence of the rapid decline in performance predicted by the between-list PI interpretation when the same set of similar items was repeated on each trial. Indeed, if anything the dissimilar conditions show more of a resemblance to this pattern than the similar, a result that is opposite to Engle’s prediction.

Figure 2 about here

Discussion

Our results appear to strongly support the assumption that the phonological similarity effect in immediate serial recall principally reflects the influence of within-list similarity. There is however evidence to suggest a modest but significant advantage to switching vowel similarity repeatedly across trials. The relative magnitude of the repetition effect in comparison to that of within-sequence similarity is small (the effect size was smaller by a factor of over 4) and emphasises the relative importance of within-sequence as against between-sequence effects. The individual trial data lend no support to the view that a build up of PI in this paradigm will resemble the dramatic effect shown in the classic Peterson task. In conclusion, while there may be evidence of a small PI effect, the evidence strongly favours the standard within-lists interpretation of the phonological similarity effect.

This leads on to the broader question of why two extensively used STM paradigms should behave so differently. The Peterson task did indeed prove to be massively dependent on the influence of PI (Keppel & Underwood, 1962; Loess, 1964), with only a small contribution to forgetting when PI was minimised by testing only a single item (Baddeley & Scott, 1971). We find the opposite pattern for immediate serial recall with little contribution from PI and a major effect from within sequence interference. It is interesting in this connection to note that when the Peterson task is combined with within-sequence phonological similarity, they do not interact. Triplets of phonologically similar words were more poorly retained than dissimilar but were forgotten at the same rate (Baddeley, 1968).

We suggest that the reason for this difference lies in the nature of the two tasks. The essence of the Peterson task is to recall the last of a sequence of items, avoiding the recall of earlier items. Typically the items themselves are well within span, three consonants or words, but when successive triplets are drawn from the same general category, performance appears to depend on some form of temporal judgement (Turvey, Brick & Osborne, 1970), reflecting the ratio of two temporal delays, that of the last item and that of the penultimate item, the so called ratio rule (Baddeley, 1986: Brown et al., 2007; Glenberg et al., 1980). As delay increases, this relative judgement becomes more difficult. When successive triplets come from readily discriminable categories, this can be used to avoid interference across trials substantially reducing dependence on temporal cues.

The essence of the immediate serial recall task on the other hand is to maintain the order of several items, typically with no delay. The cues therefore are time-based in one case and order-based in the other. Theories such as SIMPLE (Brown et al., 2001) give a good account of time-based forgetting over a wide range of delays, but have difficulty in giving an equally good account of the retention of serial order (Lewandowski, Brown, Wright & Nimmo, 2006). Similarly, models of serial order (see Hurlstone, Hitch & Baddeley, 2014 for a review) are not readily adapted to time-based forgetting.

In conclusion, although the term proactive interference provides a useful term for broadly describing the impact of early learning on later retention, its precise characteristics are likely to depend on the underlying systems and the specific tasks used to investigate them. PI appears to play an important role when retrieval depends on temporal cues to distinguish items to be remembered from earlier items, whereas the capacity to retain the serial order of a sequence of items appears to depend on within-list factors.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in Research and Theory.* (Vol. 2, pp. 89-195). New York: Academic Press.

Baddeley, A. D. (1966a). 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. (1966b). 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. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology, 20*, 249-264.

Baddeley, A. D. (1986). *Human memory*. Hove: Erlbaum.

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

Baddeley, A. D., Gathercole, S. E. & Papagno, C. !998). The phonological loop as a language learning device. Psychological Review, 105 (1), 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., & Scott, D. (1971). Short-term forgetting in the absence of proactive interference. *Quarterly Journal of Experimental Psychology, 23*(275-283).

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

Conrad, R. (1967). Interference or decay over short retention intervals? *Journal of Verbal Learning and Verbal Behavior, 6*, 49-54.

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology, 55*, 429-432.

Engle, R. W. (2007). Integrative comments Working Memory: The mind is richer than the models. In H. L. R.oediger III, Y. Dudai, & S. M. Fitzpatrick (Eds.), *Science of Memory: Concepts* (pp. 159-166). Oxford: Oxford University Press.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The Psychology of Learning and Motivation.* (Vol. 44, pp. 145-199). New York: Elsevier.

Estes, W. K. (1991). On types on item coding and sources of recall in short-term memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: In honor of Bennet B.Murdock* (pp. 175-194). Lawrence Erlbaum Associates Inc. : Hillsdale, NJ.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., Gretz, A. L., . . . Turpin, V. M. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 355-369.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140*, 339-373. doi:10.1037/a0034221.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning & Verbal Behavior, 1*, 153-161.

Lewandowsky, S., Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language, 54*, 20-38.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *Psychology of learning and Motivation: Advances in Research and Theory, 49*, 1-48.

Loess, H. (1968). Short-term memory and item similarity. *Journal of Verbal Learning & Verbal Behavior, 7*, 87-92.

Loess, H., & Waugh, N. C. (1967). Short-term memory and inter-trial interval. *Journal of Verbal Learning and Verbal Behavior, 6*, 455-460.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193-198.

Shipstead, Z., & Engle, R. W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 277-289. doi:10.1037/a0028467

Turvey, M. T., Brick, P., & Osborn, J. (1970). Proactive interference in short-term memory as a function of prior-item retention interval. *Quarterly Journal of Experimental Psychology, 22*, 142-147.

Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *American Journal of Psychology, 78*, 567-574.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 2*, 440-445.

Table 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Set A | Set B | Set C | Set D | Set E | Set F |
| man | pen | rim | cod | bud | peel |
| cat | bet | bin | cot | bug | feel |
| bat | jet | pin | fog | cub | feed |
| ban | bed | rip | hop | rub | deed |
| pad | fed | kit | log | rut | reed |
| pan | pet | pit | top | cut | feet |

Table 1. The six sets of six words used to generate the four experimental conditions, as described in the text.

Figure 1

Figure 1: Mean number of correctly recalled items in the four conditions. Error bars are standard errors of the means.

Figure 2



Figure 2: Mean number of items correctly recalled as a function of trials

1. We use the term working memory to refer to a theoretically based conceptual framework and the term short-term memory to refer to the capacity to store small amounts of material over brief periods of time, regardless of how this is interpreted theoretically. [↑](#footnote-ref-1)