

Interpreting potential markers of storage and rehearsal: Implications for studies of verbal short-term memory and neuropsychological cases

Xiaoli Wang^{1,2} · Robert H. Logie³ · Christopher Jarrold¹

Published online: 11 March 2016
© Psychonomic Society, Inc. 2016

Abstract Neuropsychological studies of verbal short-term memory have often focused on two signature effects – phonological similarity and word length – the absence of which has been taken to indicate problems in phonological storage and rehearsal respectively. In the present study we present a possible alternative reading of such data, namely that the absence of these effects can follow as a consequence of an individual’s poor level of recall. Data from a large normative sample of 251 adult participants were re-analyzed under the assumption that the size of phonological similarity and word length effects are proportional to an individual’s overall level of recall. For both manipulations, when proportionalized effects were plotted against memory span, the same function fit the data in both auditory and visual presentation conditions. Furthermore, two additional sets of single-case data were broadly comparable to those that would be expected for an individual’s level of verbal short-term memory performance albeit with some variation across tasks. These findings indicate that the absolute magnitude of phonological similarity and word length effects depends on overall levels of recall, and that these effects are necessarily eliminated at low levels of verbal short-term memory performance. This has implications for how one interprets any variation in the size of these effects, and raises serious questions about the causal direction of any relationship

between impaired verbal short-term memory and the absence of phonological similarity or word length effects.

Keywords Short-term memory · Phonological similarity · Word length

Verbal short-term memory supports our ability to recall verbal information in correct serial order, and is thought to underpin aspects of language acquisition (Baddeley, Gathercole, & Papagno, 1998). According to Baddeley’s (1986, 1992) influential model of working memory, verbal short-term memory performance is supported by the ‘phonological loop’, which itself has two components. The first of these is a phonological store that maintains information in a phonological form; the second is a rehearsal loop that offsets the forgetting from the store, which would otherwise be caused by trace decay, by reactivating memoranda through a process of subvocal re-presentation. Evidence for these two components is thought to come from two experimental phenomena that are, as a result, often taken as markers of their function. Evidence that verbal short-term memory involves phonologically-based storage comes from the phonological similarity effect (Conrad & Hull, 1964), the finding that immediate serial recall for lists of phonologically confusable items tends to be poorer than that for comparable lists of phonologically dissimilar memoranda. Support for the claim that rehearsal supports verbal short-term memory is traditionally drawn from the finding that immediate serial recall of words of a long spoken duration is poorer than that for the corresponding number of words of a shorter spoken duration – the word length effect (Baddeley, Thomson, & Buchanan, 1975). A further point to note is that, according to Baddeley’s (1986) model, auditorily presented material has obligatory access to the phonological store. In contrast, visually presented information, such as pictures of

✉ Christopher Jarrold
C.Jarrold@bristol.ac.uk

¹ School of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol BS8 1TU, UK

² Department of Psychology, Northwest Normal University, Lanzhou, China

³ Department of Psychology, University of Edinburgh, Edinburgh, UK

labelable objects, must first be recoded into a phonological form before it can be maintained within the phonological loop. This process of recoding is thought to be associated with rehearsal, because it appears to be disrupted by concurrent articulation in the same way as rehearsal processes (Baddeley, Lewis, & Vallar, 1984).

It should be noted that the status of the word length effect as a potential marker of rehearsal has been the subject of considerable debate (e.g., Beaman, Neath, & Surprenant, 2008; Brown & Hulme, 1995; Caplan, Waters, & Howard, 2012; Jalbert, Neath, Bireta, & Surprenant, 2010; Lewandowsky & Oberauer, 2008). However, in mainstream cognitive psychology authors who are sympathetic to the Baddeley model continue to interpret reductions in the word length effect or phonological similarity effect as evidence that experimental manipulations have compromised participants' ability to maintain information in verbal short-term memory, whether working with adults (Camos, Mora, & Barroulet, 2013; Lobley, Baddeley, & Gathercole, 2005) or children (Henry, Messer, Luger-Klein, & Crane, 2012; Mora & Camos, 2015; Tam, Jarrold, Baddeley, & Sabatos-DeVito, 2010).

The same holds for many studies of verbal short-term memory performance in neuropsychological patients, where the phonological similarity effect and the word length effect are often used as respective markers of the function of the phonological store and subvocal rehearsal loop (see Chiricozzi, Clausi, Molinari, & Leggio, 2008; Gorno-Tempini et al., 2008; Gvion & Friedmann, 2012; Jacquemot, Dupoux, & Bachoud-Lévi, 2011; Silveri & Baldonero, 2013; Vallat-Azouvi, Weber, Legrand, & Azouvi, 2007 for recent examples). Neuropsychological case studies that have examined word length and phonological similarity effects across auditory and visual presentation modalities have shown a variety of profiles, argued to reflect problems in either rehearsal (Vallar & Baddeley, 1984), phonological storage (Vallar, Di Betta, & Silveri, 1997), or recoding of visual information (Papagno, Lucchelli, & Vallar, 2008). However, as noted by Trojano and Grossi (1995), a particularly common pattern in studies of verbal short-term memory among neuropsychological patients is the observation of a reliable phonological similarity effect with auditory but not with visual presentation of material, and a non-significant word length effect in either modality (Belleville, Peretz, & Arguin, 1992; Bisiacchi, Cicolotti, & Denes, 1989; Howard & Franklin, 1990; Trojano & Grossi, 1995; Vallar & Baddeley, 1984; Vallar, Basso, & Bottini, 1990; Waters, Rochon, & Caplan, 1992; see also Vallar, Corno, & Basso, 1992; Vallar & Papagno, 2002). Given the arguments advanced above, this might seem to imply that an associated problem of rehearsal and recoding is a common cause of impaired verbal short-term memory. However, other work suggests alternative explanations of this pattern. In this paper we explore those alternative readings of

reduced phonological similarity and word length effects, and develop their implications for the study of verbal short-term memory in both neuropsychological patients and samples from the general population.

The starting point for our analysis is work by Logie, Della Sala, Laiacona, Chambers, and Wynn (1996), published in this journal. These authors examined the size of both the phonological similarity and the word length effects shown by a large sample ($n = 251$) of adult participants drawn from the general population. Logie et al. (1996) found that a number of their sample failed to show reliable phonological similarity or word length effects, under conditions of either auditory or visual presentation. The extent to which these effects were absent was not particularly consistent across, and even within, participants, with some individuals showing an absent effect on one measure or in one modality in combination with a reliable effect on the other measure or in the other modality. In addition, when a subsample of 40 individuals, 20 of whom failed to show one of the expected effects in the four conditions of the original experiment, was retested at a later point, very low test-retest reliability of the size of the phonological similarity or word length effects was observed. Nevertheless, at the sample level Logie et al. (1996) found that the absolute size of an individual's phonological similarity or word length effect was related to their level of overall recall, and was related to whether or not participants reported using subvocal rehearsal as a memory strategy.

Beaman et al. (2008) replicated the finding that the magnitude of the phonological similarity and word length effects shown by adult undergraduates was proportional to individuals' level of recall. Similarly, Jarrold and Citroën (2013) argued that reductions in the size of the phonological similarity effect seen among younger children (cf., Henry et al., 2012; Hitch, Halliday, Dodd, & Littler, 1989) could be explained in terms of proportional scaling; if the absolute size of the phonological similarity effect is proportional to individuals' recall capacity, then one would expect to see smaller absolute similarity effects in young children who necessarily tend to recall fewer items in tests of verbal short-term memory. In a sample of 116 children aged between 5 and 9 years of age, Jarrold and Citroën (2013) showed that age differences in absolute size of the phonological similarity effect across a variety of encoding and recall conditions were eliminated when the similarity effect was scored proportionally.

This claim for proportional scaling, echoed in other areas of cognitive psychology (e.g., Cerella, 1985), has clear and profound implications for our understanding of phonological similarity and word length effects in adults and, particularly, neuropsychological patients. Given that patients with a verbal short-term memory deficit will, almost by definition, show very low levels of immediate serial recall, and if the size of these effects is proportional to overall performance, then one would expect similarity and length effects to be smaller than

normal in such patients simply as a consequence of their memory impairment. Indeed, as Caplan et al. (2012) note, when discussing the implications of a reduced phonological similarity effect (PSE) in the context of neuropsychological studies, “lower list lengths and reduced spans are associated with a smaller PSE” (p. 294). In addition, because overall immediate serial recall performance is known to be poorer when material is presented visually as opposed to auditorily (e.g., Harvey & Beaman, 2007; Murray, 1966) one would expect this reduction to be particularly marked with visual presentation (cf. Jarrold & Citroën, 2013). If this analysis is correct, then reduced phonological similarity and word length effects among individuals with verbal short-term memory deficits, particularly with visual presentation of memoranda, would not, in and of themselves, necessarily provide any evidence of problems of rehearsal or recoding, or of reduced phonological storage capacity (cf. Caplan et al., 2012). Importantly, neuropsychological studies in this area have measured either phonological similarity or word length effects in terms of the absolute difference between span or recall scores across conditions (e.g., Belleville et al., 1992; Bisiacchi et al., 1989; Chiricozzi et al., 2008; Gorno-Tempini et al., 2008; Gvion & Friedmann, 2012; Howard & Franklin, 1990; Jacquemot et al., 2011; Papagno et al., 2008; Silveri & Baldonero, 2013; Trojano & Grossi, 1995; Vallar & Baddeley, 1984; Vallar et al., 1990; Vallar et al., 1992; Vallar et al., 1997; Vallat-Azouvi et al., 2007; Waters et al., 1992). Indeed, to our knowledge no neuropsychological report of short-term memory function, including those published since Logie et al.’s (1996) paper, has employed proportional scoring of these effects.

To test the suggestion that phonological similarity and word length effects scale proportionally in adult samples, the current paper re-analyzes data from two previous studies. We first re-analyzed Logie et al.’s (1996) data to test this prediction, and to see whether a proportional scaling account holds equally across different presentation conditions. Our second set of analyses took advantage of the fact that Logie and colleagues excluded from their 1996 paper one participant who showed, in their view, atypically small phonological similarity and word length effects and who themselves complained of suffering from memory problems. The data from that participant were subsequently published by Della Sala and Logie (1997), who noted that this individual had experienced a severe case of chicken pox early in childhood that had delayed the onset of their formal education. Della Sala and Logie were careful to make clear that they had no direct evidence for any resultant neurological impairment in this individual, but suggested that their apparently atypical performance may well have reflected “some form of brain damage resulting in impaired function of the phonological loop” (p. 380). Here we examine the extent to which the size of this individual’s phonological similarity and word length effects were indeed

atypical, as Della Sala and Logie (1997) suggested, given their overall levels of verbal short-term memory performance. Finally, we also explore the performance of a further subject from the original dataset reported by Logie et al. (1996) who had relatively high span scores but small phonological similarity and word length effects under certain presentation conditions.

Method

Full procedural details for the experiment that provided the data that are re-analyzed here are available in Logie et al. (1996). The main participants ($n = 251$, 117 male) were drawn from the general population and were aged between 18 and 70 years ($M = 42.8$, $SD = 3.1$). The additional “Subject 236” was drawn from the same population, was male, and was aged 36 years at the time of testing (Della Sala & Logie, 1997). “Subject 37” (see below) formed part of the original data set reported by Logie et al. (1996).

Four stimulus sets were employed, each containing nine items. Two contained phonologically dissimilar or phonologically similar words, respectively, matched for frequency, in order to evaluate the phonological similarity effect. The remaining two contained words of a short- or long-spoken duration, respectively, again matched for frequency, to examine the word length effect. Each stimulus set was presented in two span tasks; in one the memoranda were presented auditorily by the experimenter, in the other stimuli were presented visually as written words. In both cases stimuli were presented at a rate of one item per second, and recall was oral. The span procedure employed involved presentation of three trials at each list length, starting at list length 2, with participants moving on to the next list length if they correctly recalled all items on two successive occasions. The dependent variable taken from each task was the mean list length of the last three sequences that the participant correctly recalled. Here we report two different indices of the phonological similarity and word length effects. The first is the absolute size of these effects (dissimilar score – similar score; short word score – long word score) and the second is the proportional effect size (absolute PSE/dissimilar score; absolute WLE/short word score) (cf. Beaman et al., 2008).

Results

The phonological similarity effect in the main sample

Figure 1 reproduces part of Fig. 1 from Logie et al.’s (1996) paper, and shows the mean span score performance for the four conditions involved in the assessment of phonological similarity. An initial analysis, not reported in the original

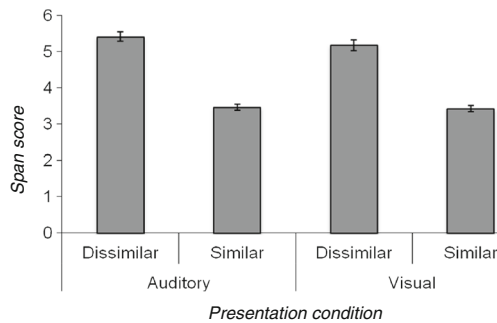


Fig. 1 Recall performance of the main sample for phonologically dissimilar and phonologically similar words (error bars are 95 % confidence intervals)

paper, confirmed that immediate recall of just dissimilar items was superior with auditory as opposed to visual presentation, $F(1, 250) = 10.974$, $p < .001$, $MSE = 0.605$, $\eta_p^2 = .042$. This confirms the presence of the expected modality effect in these data. Logie et al. (1996) reported details of the absolute size of the phonological similarity effect in both presentation conditions, noting that it was significant for both auditory, $F(1, 250) = 890.684$, $p < .001$, $MSE = 0.534$, $\eta_p^2 = .781$, and visual, $F(1, 250) = 532.376$, $p < .001$, $MSE = 0.724$, $\eta_p^2 = .680$, presentation. Here we additionally note that the absolute phonological similarity effect was significantly larger under auditory than under visual presentation, $F(1, 250) = 5.099$, $p = .025$, $MSE = 0.915$, $\eta_p^2 = .025$. In addition, the correlation between the size of each individual's phonological similarity effect across the two presentation conditions was significant, $r(250) = .278$, $p < .001$.

The fact that the phonological similarity effect was significantly smaller in the presentation condition that was associated with generally lower levels of recall is clearly consistent with the suggestion that the absolute size of the effect is proportional to level of performance. To investigate this possibility further, Fig. 2 plots the proportional phonological similarity effect shown by each individual against their level of dissimilar recall, for either the auditory (panel a) or visual (panel b) presentation conditions (note, similar plots, albeit without showing individual data, are given in Fig. 2 of Logie et al., 1996). One might intuitively expect these functions to be entirely flat if the size of the effect is proportional to level of performance, and this is clearly not what is observed. However, a similar non-linear developmental pattern was also seen in Jarrold and Citroën's (2013) data. Jarrold and Citroën found that phonological similarity effects were proportional to children's "baseline" levels of recall (as indexed by their recall of phonologically dissimilar lists) with the exception of one cell of their design, where the youngest children performed the hardest version of the task (visual presentation and verbal recall). Here phonological similarity effects were smaller than predicted even when coded proportionally, and were, on occasions, negative. Subsequent work by Jarrold, Danielsson,

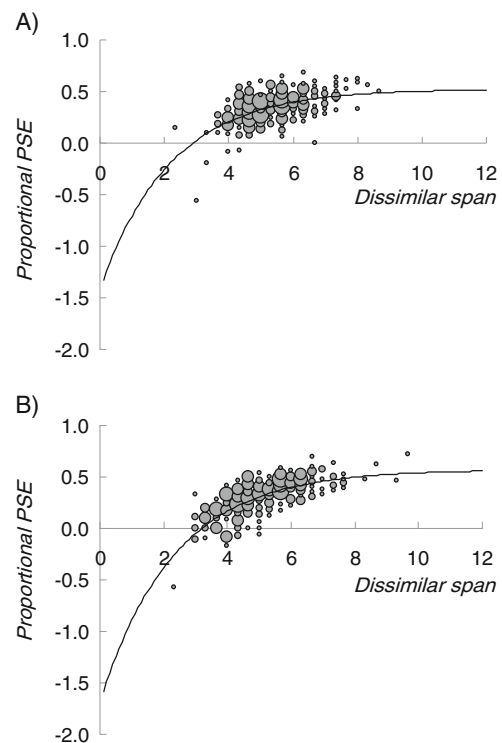


Fig. 2 Plots of each individual's proportional phonological similarity effect against their level of dissimilar recall for auditory presentation (panel a) and visual presentation (panel b) conditions. Multiple data points are shown by corresponding larger circles

and Wang (2015) confirmed that the Jarrold and Citroën data could be modeled using a negative exponential growth function. This simultaneously captures the fact that proportionalized phonological similarity effects are small and sometimes negative at low levels of recall, plus the finding that proportional effects of similarity tend to a constant value across higher levels of performance.

Further simulations by Jarrold et al. (2015) showed that randomly-generated datasets based on proportional scaling of any manipulation effect, combined with noise and the assumption of a floor to possible levels of recall, produced similar functions. Specifically, when overall levels of recall are low, proportional scaling predicts a small absolute effect of any manipulation, and noise in the estimate of the "harder" condition can outweigh this difference producing a negative manipulation effect on occasions. In addition, proportionalizing scores by dividing the absolute manipulation effect by performance in the easier (e.g., dissimilar) condition can lead to large negative proportional values, particularly when recall in this easier condition is low. However, positive manipulation effects can never exceed 1 when proportionalized, because this represents the limit obtained when recall in the harder condition is zero. As a result, the consequence of noise associated with the measurement of these values (possibly due to variability in strategy use), coupled with the necessary presence of floor effects in any assessment, is that a plot of proportional effect

size against level of recall (in this case dissimilar recall) will be a curve that begins with large negative values but which asymptotes at the constant level of proportional cost associated with the effect. Again, Jarrold et al. (2015) showed that negative exponential growth curves provided a satisfactory fit to such functions.

Figure 2 clearly shows that the functions derived from the Logie et al. (1996) data also take this form; indeed, the two datasets show a very similar pattern. To examine the degree of comparability of the two panels of Fig. 2, and following Jarrold et al. (2015), negative exponential growth functions were fit to each dataset with three free parameters in each case (an intercept, an asymptote, and a growth rate parameter). The fit values and parameter estimates for these models are shown in Table 1, which shows that these functions provided reliable fits to the data in each case. Table 1 also shows two more measures that can be used to compare the models, namely the final asymptote value and the estimated point of X-axis intercept.

Although the precise values of these indices differ across the two models, they are reasonably similar, suggesting that in fact a single negative exponential growth function might fit both sets of data simultaneously. To that end we then compared the relative fits of a single three-parameter model through both datasets with that of a model that allowed the three free parameters for each dataset to vary independently of each other (a six-parameter model). Although the six-parameter model explained somewhat more variance than the simpler model, a formal comparison of the goodness of fit of these two models showed no significant difference between them (SS for three-parameter model = 59.757, SS for six-parameter model = 59.839, $F = .001$, $p > .999$). On the basis of parsimony we therefore prefer the simpler model, which implies that the proportional phonological similarity effect is directly comparable across the two presentation conditions of Logie et al.'s (1996) experiment. In other words, although visual presentation led to a significantly smaller absolute effect of similarity than auditory presentation, this can be entirely accounted for statistically by the fact that baseline levels of recall were smaller in the visual presentation condition as shown above. This leaves open the question of whether the smaller visual presentation effect reflects the reduced ability of visual coding to retain serial verbal order (e.g.,

Logie, Saito, Morita, Varma, & Norris, 2015; Saito, Logie, Morita & Law, 2008) or impaired phonological loop functioning.

The word length effect in the main sample

Figure 3 again reproduces a section of Logie et al.'s (1996) Fig. 1, this time plotting the span score performance for the four tasks relevant to the examination of the word length effect in their dataset. Once again an analysis of only short word performance revealed a significant modality effect, $F(1, 250) = 28.886$, $p < .001$, $MSE = 0.406$, $\eta_p^2 = .104$, due to superior recall with auditory presentation. As noted by Logie et al. (1996), the absolute word length effect was significant for both auditory, $F(1, 250) = 277.421$, $p < .001$, $MSE = 0.507$, $\eta_p^2 = .526$, and visual, $F(1, 250) = 194.833$, $p < .001$, $MSE = 0.539$, $\eta_p^2 = .438$, presentation. However, here we additionally report that the absolute size of the word length effect was significantly larger with auditory than with visual presentation, $F(1, 250) = 4.346$, $p < .001$, $MSE = 0.596$, $\eta_p^2 = .017$. A further analysis showed a significant correlation between the size of each individual's word length effect across the two presentation modalities, $r(250) = .436$, $p < .001$.

Figure 4 plots the proportional word length effect shown by each individual in the main sample against their level of short word recall, for both auditory and visual presentation. As with the similarity effect, negative exponential growth curves provided good fits to each dataset (see Table 2). Again the two models produced comparable parameter estimates, suggesting that a single negative exponential function might be applied to both datasets. We therefore compared a single three-parameter model of both the auditory and visual data combined with a six-parameter model that allowed for different intercept, asymptote, and rate values for each presentation mode. While the latter model accounted for more variance in the data, a formal comparison of their goodness of fit provided no evidence to support the more specified model, (SS for three-parameter model = 29.619, SS for six-parameter model = 29.753, $F < .001$, $p > .999$). We therefore again prefer the simpler model which indicates that word length effects scale proportionally in the same way under both auditory and visual presentation conditions in this large sample, and that the greater absolute word length effect under auditory presentation is

Table 1 Negative exponential growth function fits to Fig. 2. Intercept, asymptote, and rate are the three parameters of each model: similar span = (Int. + Asym. * (1-EXP(-Rate * dissimilar word span)))

Presentation mode	Intercept	Asymptote	Rate	R ²	Final asymptote	X-axis intercept
Auditory	-1.366	1.890	0.454	.292	0.52	2.90
Visual	-1.626	2.204	0.430	.531	0.57	3.20

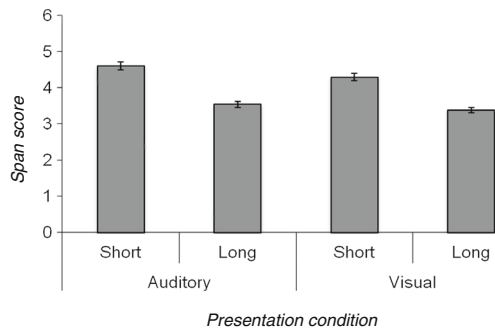


Fig. 3 Recall performance of the main sample for short and long words (error bars are 95 % confidence intervals)

simply a consequence of this mode giving rise to higher baseline levels of performance.

Subject 236

Figure 5 plots the performance of the single individual reported separately by Della Sala and Logie (1997), and does so by superimposing their proportional phonological similarity effects and word length effects on the normative curves in Figs. 2 and 4. This individual was tested twice on all tasks, so Fig. 5 plots both their initial and re-test phonological similarity effects (A - auditory presentation; B - visual

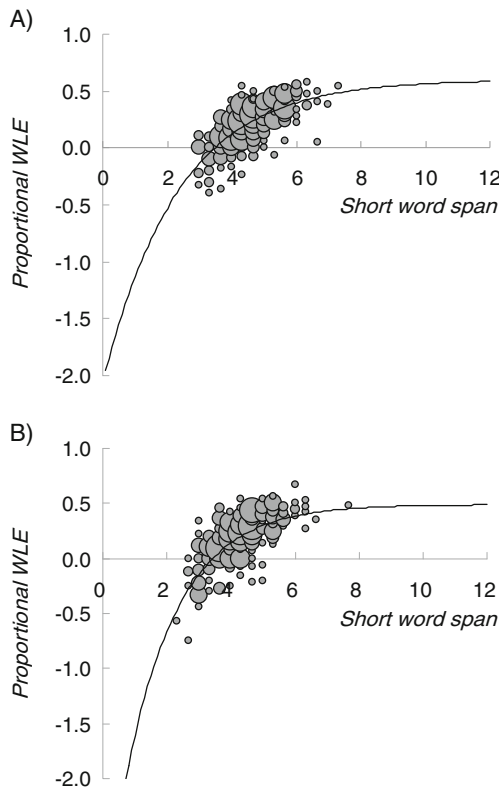


Fig. 4 Plots of each individual’s proportional word length effect against their level of short word recall for auditory presentation (panel a) and visual presentation (panel b) conditions. Multiple data points are shown by corresponding larger circles

Table 2 Negative exponential growth function fits to Fig. 4. Intercept, asymptote, and rate are the three parameters of each model: long word span = (Int. + Asym. * (1-EXP(-Rate * short word span))

Presentation mode	Intercept	Asymptote	Rate	R ²	Final asymptote	X-axis intercept
Auditory	-2.014	2.610	0.426	.481	0.60	3.40
Visual	-3.226	3.724	0.604	.493	0.49	3.40

presentation) and word length effects (C - auditory presentation; D - visual presentation). The figure clearly shows that this individual has impoverished verbal short-term memory performance as their span score is close to floor (a span of 2 on all conditions, particularly at the time of initial testing). Nevertheless, while they do indeed show small or absent similarity and length effects, as Fig. 5 illustrates, in seven out of eight cases, the proportional phonological similarity or word length effect is numerically higher than would be predicted from the normative function, and there is no evidence in these data that these effects are any smaller than one would expect given the typical pattern of proportional costs. This is particularly striking for the second test session on word length with visual presentation, indicating that the participant may be attempting subvocal rehearsal, even if their verbal STM is low. This reinforces the point that a small manipulation effect cannot necessarily be interpreted as suggesting that a participant with a poor verbal span is incapable of using subvocal rehearsal and phonological storage.

Residual variance and Subject 37

Finally, when fitting functions to aggregate data, as noted earlier, it is common to consider the residual variance as noise due to measurement error. However, a major implication of the Logie et al. (1996) paper is that some of this apparent measurement error may be explained by participants not performing a task in the way that the experimenter expects. Logie et al. (1996) noted that participants reported sometimes using subvocal rehearsal, sometimes using visual imagery, sometimes remembering the first letter of a word, and sometimes using other strategies. This variability in strategy would generate what looks like measurement error when averaging results across all participants. However, the demonstration by Logie et al. (1996) that reported strategy accounted for the magnitude of the four phenomena of interest, even when individual span was taken into account, indicates that not all of the residual variance is random noise. So, the combination of proportional scaling and variation in strategy use by participants may contribute to the lack of the standard effects, reinforcing the point that the lack of these effects does not necessarily indicate an inability to use subvocal rehearsal or phonological storage. More recently,

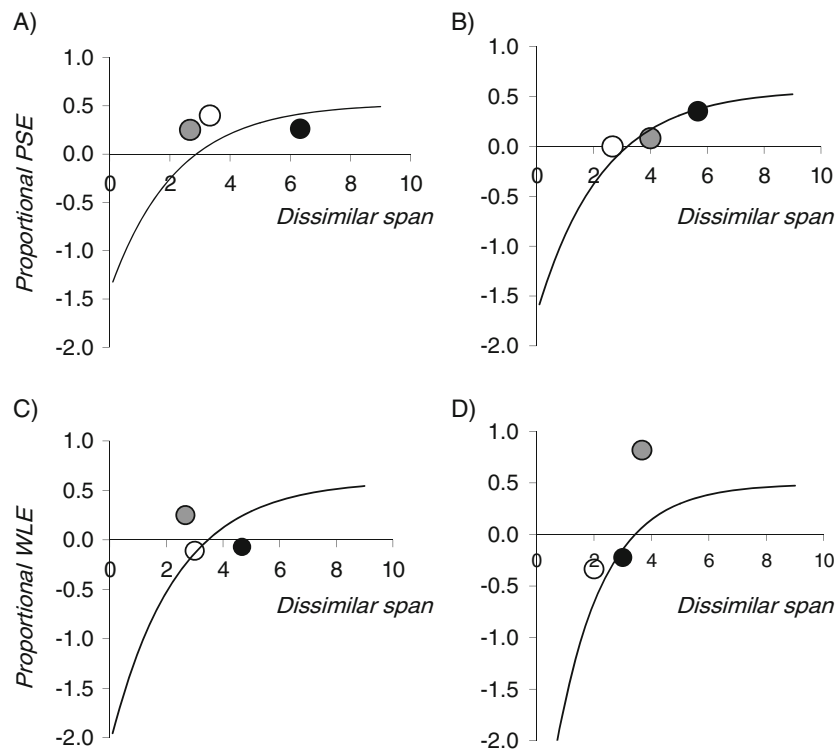


Fig. 5 The proportional phonological similarity effects (with auditory presentation in panel a and visual presentation in panel b) and word length effects (with auditory presentation in panel c and visual

presentation in panel d) shown by Subject 236 and Subject 37. Empty circles are first assessment point for Subject 236, grey circles are second assessment point for Subject 236, black circles are data for Subject 37

Johnson, Logie, and Brockmole (2010) demonstrated the importance of examining task-specific residual variance when exploring common variance across a battery of tests, suggesting that different participants may perform the same tasks in different ways. Their results further suggested that participants who perform poorly may be using strategies that are ineffective for supporting performance on a given task, such as attempting to use verbal coding to remember abstract visual patterns. In this context it is interesting to consider case number 37 included in the Logie et al. (1996) analyses, whose data are plotted in proportional terms in Fig. 5. This participant had relatively high span scores yet showed better proportionally scaled performance with long words than with short words for both visual (-0.233) and auditory (-0.071) presentation. The discrepancy from the function relating span to effect magnitudes is most clear in the case of auditory presentation (Fig. 5c). In the absence of an alternative account this might be interpreted as measurement noise, but as demonstrated by Logie et al. (1996) and Johnson et al. (2010), measurement variability may instead reflect, at least in part, this participant using a strategy other than subvocal rehearsal to remember the long and short words. We no longer have the original raw data to allow us to link this particular participant with a reported strategy. However, the possibility that this intra-task variation in proportionalized effect sizes is due to variation in strategy use is wholly consistent with the report by Della Sala, Logie,

Marchetti, and Wynn (1991) of a single case participant with a digit span of 9 who consistently failed to show the standard effects until instructed to use subvocal rehearsal.

Discussion

The aim of the present study was to explore what one can infer from cases of reduced phonological similarity and word length effects in adults, particularly adults who, for whatever reason, perform poorly on tests of verbal short-term memory. This work is therefore relevant to neuropsychological studies of short-term memory, not least because a common assumption in such studies is that a lack of these effects is a marker of a failure in the function of the phonological store and rehearsal subcomponents of Baddeley's (1986, 1992) phonological loop model (e.g., Belleville et al., 1992; Jacquemot et al., 2011; Papagno et al., 2008; Silveri & Baldonero, 2013; Vallar & Baddeley, 1984; Vallar et al., 1997). However, it is also directly applicable to child and adult studies of verbal short-term memory in the general population, where participants do vary in the extent to which they show phonological similarity and word length effects.

Indeed, our first analysis re-examined the variation in the size of these effects within a sample from the general population first reported by Logie et al. (1996). As shown in Fig. 2,

for both auditory (A) and visual (B) presentation conditions, when the proportional phonological similarity effect shown by each individual in this dataset was plotted against their dissimilar memory span, a negative exponential growth function fitted the data well. An entirely similar pattern was observed when individuals' word length effects were plotted against their long word span (see Fig. 4). These curves are consistent with existing data from developmental populations (Jarrold, 2013; Jarrold & Citroën, 2013) that also can be fit with this form of negative exponential growth function (Jarrold et al., 2015).

We do not wish to claim that a negative exponential growth function necessarily represents the only way one might model these data. The two aspects of the data that any function would need to capture would be (i) the evidence that proportionalized scores "level off" to a fixed asymptotic value, reflecting what we believe to be the essentially proportional nature of the costs of similarity and length (cf. Beaman et al., 2008), and (ii) the presence in the data of negative effects when overall performance levels are low. A negative exponential growth function captures these aspects of the data in its asymptote and intercept parameters, respectively.

It is important to stress that negative phonological similarity and word length effects do not arise *because* we have proportionalized these indices; they are already present in the data whether measured in absolute or proportional terms. Rather, the effect of proportionalizing these effects is simply to exacerbate the influence of these negative values on the resultant function. This is a potential concern with the proportionalizing approach, but one that we argue is offset by the benefit of being able to fit a function that asymptotes to an index of the fixed proportional cost of the manipulation. Indeed, the models summarized in Tables 1 and 2 do provide a significant fit to the data, and so capture meaningful variance in performance even given the fact that low levels of recall are often associated with negative effects of similarity or of length.

It is possible that these negative manipulation effects, which we believe can arise when overall performance levels are low, would not be seen to the same extent in studies that employ more sensitive measures of recall, and indeed the majority of neuropsychological studies present considerably more trials to measure recall performance at a given list length than the three employed by Logie et al. (1996) (e.g., Belleville et al., 1992; Bisiacchi et al., 1989; Gorno-Tempini et al., 2008; Gvion & Friedmann, 2012; Jacquemot et al., 2011; Papagno et al., 2008; Silveri & Baldonero, 2013; Trojano & Grossi, 1995; Vallar & Baddeley, 1984; Vallar et al., 1990; Vallar et al., 1997; Waters et al., 1992; though see by contrast Chiricozzi et al., 2008; Vallar et al., 1992; Vallat-Azouvi et al., 2007). However, the significant correlations between the size of individuals' similarity and word length effects across the two presentation modalities (auditory vs. visual), reported for the first time in the current paper, indicates that

these effects were not entirely unreliable in the Logie et al. (1996) data. In addition, even if tested under conditions that led to higher reliability, which might well reduce the number of negative effects observed, the proportional scaling account would still predict smaller, albeit positive, effects of similarity and length at low levels of recall when measured in absolute terms. This could result in both a non-significant effect of the manipulation amongst such individuals, and a significant interaction between the absolute size of the manipulation and group if such a sample were compared to individuals with higher overall spans (cf. Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2011).

It is also worth noting that negative phonological similarity and word length effects have been reported in other studies of verbal short-term memory (Campoy & Baddeley, 2008; Carlesimo, Galloni, Bonanni, & Sabbadini, 2006; Copeland, & Radvansky, 2001; Fallon, Groves, & Tehan, 1999; Henry, Turner, Smith, & Leather, 2000; Lian, & Karlsen, 2004; Romani, McAlpine, Olson, Tsouknida, & Martin, 2005), and, crucially, often occur in the general adult population in conditions that increase task difficulty relative to traditional immediate serial recall methods (Copeland & Radvansky, 2001; Fallon et al., 1999; Lian & Karlsen, 2004; Romani et al., 2005). This may be because under such conditions participants adopt a strategy other than rehearsal, such as using the similarity between phonologically similar items as a cue to item identity (Fallon et al., 1999; Gupta, Lipinski, & Aktunc, 2005), or making use of semantic information that might be richer for long than for short words (Campoy & Baddeley, 2008). However, detection of the use of alternative strategies can only rely on positive evidence for the use of such strategies, such as effects of visual similarity (e.g., Logie et al., 2015; Saito et al., 2008) or semantic similarity (e.g., Baddeley, 1966; Campoy & Baddeley, 2008). Therefore, without such positive evidence for alternative strategies, the absence of a phonological similarity or word length effect does not imply that phonological coding and rehearsal are not being used. Nevertheless, when levels of overall recall performance are low, the absolute cost of the manipulation of phonological similarity or word length predicted by proportional scaling is small. Consequently, the impact of noise in the estimate of the manipulation effect will be more noticeable, and on occasions lead to a reversal of the predicted effect. This tendency will, in turn, be exacerbated by floor effects. Floor effects will particularly constrain the estimate of the condition that is expected to be more difficult (the phonologically similar condition, the long word condition), and mean that noise is more likely to raise rather than reduce this estimated value.

The first key implication of the current work, therefore, is that the data from this large normative sample provide support for the view that phonological similarity and word length effects scale proportionally in adults. This is shown by the fact

that the size of these effects is proportional to level of performance, once performance reaches the point at which the influence of floor effects is no longer apparent. Although the parameters of each of the functions shown in Figs. 2 and 4 are different, the deduced final asymptote values, which represent the underlying proportional cost of the manipulation, are similar, falling around .5 (see Tables 1 and 2). A 50 % cost of phonological similarity or of word length is somewhat larger than the corresponding value reported in previous studies that have measured these effects in proportional terms (Beaman et al., 2008; Logie et al., 1996). One reason for this is that some studies (e.g., Beaman et al., 2008) presented fixed list lengths of to-be-remembered items to participants. This, in contrast to the span procedure employed for the current data, can on occasion lead to ceiling effects in recall that would necessarily reduce the size of any manipulation effect. More importantly, final asymptote values in the present study are deduced from the nonlinear regression function, not by averaging each individual's effect size. The negative effects shown by individuals at low levels of recall (see Figs. 2 and 4) will necessarily reduce the overall effect size when all values are simply averaged.

A second, key point that follows from our analysis is that the effects of presentation modality on the absolute size of the phonological similarity and word length effects can be explained statistically in terms of proportional scaling, and without necessarily assuming that either condition makes more demands on recoding or rehearsal processes. Our initial analysis of the dataset confirmed the presence of a modality effect on the absolute size of both the phonological similarity effect and the word length effect, with greater absolute effects seen under auditory presentation conditions (see Figs. 1 and 3). However, baseline levels of performance were superior under auditory than visual presentation as would be expected. Indeed, when these effects were scored in proportional terms a single negative exponential growth function provided a good fit to both the visual and auditory datasets for both the phonological similarity effect (Fig. 2) and the word length effect (Fig. 4). In other words, when the beneficial effect of auditory presentation is accounted for, and these manipulation effects are scored in proportional terms, the effect of presentation modality on the phonological similarity and word length effects disappears.

An important implication of this finding is that it indicates that the reduction in the absolute size of these effects with visual, relative to auditory, presentation cannot necessarily be taken as evidence of the cost or difficulty of recoding or rehearsal of visually presented material. By extension, any other manipulation that also increases overall task difficulty will necessarily also lead to a reduction in the absolute size of phonological similarity or word length effects. For example, articulatory suppression is assumed to block rehearsal, and previous studies have shown that the word length effect persists under conditions of suppression when material is

presented auditorily, but not when it is presented visually (Baddeley et al., 1975). However, this apparently specific effect of articulatory suppression could, instead, reflect the fact that absolute levels of recall are lower with visual as opposed to auditory presentation of material, and are therefore reduced to a particularly low level by the imposition of concurrent suppression (cf. Jones, Macken, & Nicholls, 2004).

The fact that the effect of modality on the size of the phonological similarity effect and word length effect can be explained by proportional scaling also has very important implications for the study of these manipulations in neuropsychological patients with poor verbal short-term memory. As already highlighted, many studies have shown that such individuals are more likely to show phonological similarity and word length effects (scored in absolute terms) when material is presented auditorily as opposed to visually. However, the current data suggest that this difference is a simple consequence of the fact that these effects are always smaller in absolute terms when material is presented visually. An absence of these effects for visual but not auditory presentation in neuropsychological patients with poor immediate serial recall performance does not, therefore, necessarily provide good evidence that these individuals are unable to recode visual information into a phonological form. Similarly, given that the phonological similarity effect is associated with a larger experimental effect size than the word length effect in adults (Logie et al., 1996), one would expect the phonological similarity effect with auditory presentation to be the last "effect" to be reduced to a non-significant level by proportional scaling among individuals with reduced spans. A pattern of data in which patients show non-significant word length effects in either modality and a significant phonological similarity effect in only the auditory modality therefore provides no strong evidence for a combined problem of recoding and rehearsal, contrary to what has often been suggested in the neuropsychological literature (Trojano & Grossi, 1995; Waters et al., 1992; see Caplan et al., 2012).

This position is strongly supported by the data from participant 236, who produced a pattern of data similar to that shown by many single case study patients with verbal short-term memory deficits. Della Sala and Logie (1997) noted that, when first tested, this participant only showed evidence of a measurable phonological similarity effect under auditory presentation, and no evidence of a word length effect in either modality. However, when these effect sizes are plotted proportionally along with those from this individual's second testing session (see Fig. 5), there is no evidence that they deviate in any meaningful way from the typical pattern. Although some of these effects clearly are very small or even negative (i.e., reversed phonological similarity or word length effects), most of them are higher than the value predicted by the normative function, with seven of eight of these values falling above the

typical curve. In other words, these proportional effects sizes are in line with the typical trend of phonological similarity and word length effects, and are therefore comparable to those that would be expected for an individual with this general level of verbal short-term memory performance. As a result, the absence of absolute phonological similarity or word length effects in this participant's data may reflect a substantially reduced immediate memory capacity as indicated by their low span, but does not provide positive evidence that their phonological store or rehearsal mechanism functions in a qualitatively different way from that of other individuals in the sample with equally poor verbal short-term memory performance.

Similarly, the overall effect of presentation modality, observed in the main analysis of the Logie et al. (1996) data presented above, could well reflect the fact that visually presented material can potentially be encoded and retained in a non phonological form, which might be less effective than phonological coding for retaining serial order (Logie, Della Sala, Wynn & Baddeley, 2000; Logie et al., 2015; Saito et al., 2008). Warrington and Shallice (1972) reported that the brain-damaged patient, KF, only showed a severe verbal short-term memory deficit with auditory presentation. With visual presentation KF's verbal memory span was within the normal range for healthy participants when they attempt to remember visually presented material while subvocal rehearsal is prevented by articulatory suppression. If healthy participants sometimes rely on a visual code to maintain visually presented verbal material, then their lower memory span could reflect the use of this inefficient coding for serial verbal recall. As a result if a participant has a low span for visually presented material then one cannot infer that they are entirely unable to use subvocal rehearsal or phonological coding.

Our arguments here are consistent with Logie et al.'s (1996) original conclusion that individual differences in participants' recall performance could be related to variation in strategy usage, which would in turn affect the absolute size of both phonological similarity and word length effects due to proportional scaling; if participants adopt a strategy that leads to relatively poor overall recall, these effects will be reduced. Logie et al. (1996) noted that similarity and length effects were absent for participants who reported using strategies other than subvocal rehearsal, for example, remembering only the first letter of words presented, using visual imagery, or relying on semantic strategies. Logie et al. (1996) therefore argued that a substantial minority of participants were not using subvocal rehearsal to retain the word sequences, or were not doing so consistently from trial to trial or across different testing sessions. In a previous study, Della Sala et al. (1991) demonstrated that a participant with a very high span consistently failed to show the four effects because he was using imagery mnemonics and only showed the effects when he was specifically instructed to use subvocal rehearsal to remember the words.

Logie and colleagues (Della Sala & Logie, 1997; Logie et al., 1996) therefore cautioned against using a single observation of an absent phonological similarity or word length effect as evidence that an individual is incapable of phonological storage or rehearsal. They also argued that subvocal rehearsal is an optional, not an obligatory, strategy for immediate verbal serial recall, and so a failure to observe word length or phonological similarity effects in healthy adults is not evidence against the phonological loop hypothesis. Rather, this is evidence for the possibility that participants are not consistently using the options of subvocal rehearsal and phonological coding to perform immediate verbal serial recall tasks. This variability in strategy use might therefore be a major source of variability in data patterns that might otherwise be interpreted as measurement noise. Indeed, the data from Subject 37 and from Della Sala et al. (1991) suggest that the lack of these effects in participants with higher spans cannot necessarily be taken as evidence that the verbal short-term memory system functions in a qualitatively different way from that of other individuals with equally good task performance. In this case, and as noted by Logie et al. (1996), the participants may simply choose not to use subvocal rehearsal to perform the task, and instead rely on alternative strategies (cf. Johnson et al., 2010; Logie et al., 2000, 2015; Saito et al., 2008).

These observations point to the importance of developing models of the cognitive resources that participants have available to perform tasks, rather than models of individual tasks, such as immediate serial recall, that can be performed in a range of different ways using different combinations of cognitive resources (see Logie, 2011; *in press* for more detailed discussions). They also caution against assuming that proportional scaling of phonological similarity and word length effects necessarily undermines the phonological loop model of verbal short-term memory. We readily accept that individuals suffering from a specific (or indeed general) impairment of verbal short-term memory may have a reduced capacity phonological store or rehearsal difficulties. However, our key point is to argue that this cannot be inferred from the absence of measurable phonological similarity and word length effects (cf. Caplan et al., 2012). The novel advance that we are able to make here is to show that reduced verbal short-term memory capacity will itself necessarily lead to reduced, absent, or even reversed phonological similarity and word length effects due to proportional scaling coupled with floor effects. In addition, and as already noted, the combination of measurable phonological similarity or word length effects with auditory presentation and absent effects with visual presentation does not necessarily implicate a problem of recoding, but could simply reflect the fact that individuals with low verbal short-term memory spans are more likely to perform above floor with auditory than visual presentation. Finally, individuals who show high levels of span performance but small or absent

phonological similarity and word length effects may do so by virtue of their use of alternative memory strategies. Future studies of verbal short-term memory should therefore be extremely cautious when making any theoretical claims based on the absence of these effects, or their moderation by presentation modality. Rather, work in this area needs to acknowledge that the absence of such effects will necessarily follow as a consequence of individuals having poor overall performance on specific tasks that require immediate serial ordered verbal recall, and further understanding in this area is most likely to arise from detailed exploration and modeling of the range of underlying cognitive functions that participants bring to bear when attempting to perform those tasks.

Acknowledgments This research was supported by a China Scholarship Council fellowship to Xiaoli Wang.

References

- Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic, and formal similarity. *The Quarterly Journal of Experimental Psychology*, *18*, 362–365. doi:10.1080/14640746608400055
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. (1992). Working memory. *Science*, *255*, 556–559. doi:10.1126/science.1736359
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173. doi:10.1037/0033-295X.105.1.158
- Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, *36A*, 233–252. doi:10.1080/14640748408402157
- Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, *14*, 575–589. doi:10.1016/S0022-5371(75)80045-4
- Beaman, C. P., Neath, I., & Surprenant, A. M. (2008). Modeling distributions of immediate memory effects: No strategies needed? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 219–229. doi:10.1037/0278-7393.34.1.219
- Belleville, S., Peretz, I., & Arguin, M. (1992). Contribution of articulatory rehearsal to short-term memory: Evidence from a case of selective disruption. *Brain and Language*, *43*, 713–746. doi:10.1016/0093-934X(92)90092-S
- Bisiacchi, P. S., Cipolotti, L., & Denes, G. (1989). Impairment in processing meaningless verbal material in several modalities: The relationship between short-term memory and phonological skills. *The Quarterly Journal of Experimental Psychology*, *41A*, 293–319. doi:10.1080/14640748908402367
- Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed. *Journal of Memory and Language*, *34*, 594–621. doi:10.1006/jmla.1995.1027
- Camos, V., Mora, G., & Barrouillet, P. (2013). Phonological similarity effect in complex span task. *The Quarterly Journal of Experimental Psychology*, *66*, 1927–1950. doi:10.1080/17470218.2013.768275
- Campoy, G., & Baddeley, A. (2008). Phonological and semantic strategies in immediate serial recall. *Memory*, *16*, 329–340. doi:10.1080/09658210701867302
- Caplan, D., Waters, G., & Howard, D. (2012). Slave systems in verbal short-term memory. *Aphasiology*, *26*, 279–316. doi:10.1080/02687038.2011.642795
- Carlesimo, G. A., Galloni, F., Bonanni, R., & Sabbadini, M. (2006). Verbal short-term memory in individuals with congenital articulatory disorders: New empirical data and review of the literature. *Journal of Intellectual Disability Research*, *50*, 81–91. doi:10.1111/j.1365-2788.2005.00725.x
- Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin*, *98*, 67–83. doi:10.1037/0033-2909.98.1.67
- Chiricozzi, F. R., Clausi, S., Molinari, M., & Leggio, M. G. (2008). Phonological short-term store impairment after cerebellar lesion: A single case study. *Neuropsychologia*, *46*, 1940–1953. doi:10.1016/j.neuropsychologia.2008.01.024
- Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion, and memory span. *British Journal of Psychology*, *55*, 429–432. doi:10.1111/j.2044-8295.1964.tb00928.x
- Copeland, D. E., & Radvansky, G. A. (2001). Phonological similarity in working memory. *Memory and Cognition*, *29*, 774–776. doi:10.3758/BF03200480
- Della Sala, S., & Logie, R. H. (1997). Impairments of methodology and theory in cognitive neuropsychology: A case for rehabilitation? *Neuropsychological Rehabilitation*, *7*, 367–385. doi:10.1080/1713755541
- Della Sala, S., Logie, R. H., Marchetti, C., & Wynn, V. (1991). Case studies in working memory: A case for single cases? *Cortex*, *27*, 169–191. doi:10.1016/S0010-9452(13)80122-0
- Fallon, A. B., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When CAT helps RAT, but not MAN. *International Journal of Psychology*, *34*, 301–307. doi:10.1080/002075999399602
- Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., ... Miller, B. L. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology*, *71*, 1227–1234. doi:10.1212/01.wnl.0000320506.79811.da
- Gupta, P., Lipinski, J., & Aktunc, E. (2005). Reexamining the phonological similarity effect in immediate serial recall: The roles of type of similarity, category cuing, and item recall. *Memory and Cognition*, *33*, 1001–1006. doi:10.3758/BF03193208
- Gvion, A., & Friedmann, N. (2012). Phonological short-term memory in conduction aphasia. *Aphasiology*, *26*, 579–614. doi:10.1080/02687038.2011.643759
- Harvey, A. J., & Beaman, P. C. (2007). Input and output modality effects in immediate serial recall. *Memory*, *15*, 693–700. doi:10.1080/09658210701644677
- Henry, L. A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual, and semantic coding strategies and children's short-term picture memory span. *The Quarterly Journal of Experimental Psychology*, *65*, 2033–2053. doi:10.1080/17470218.2012.672997
- Henry, L. A., Turner, J. E., Smith, P. T., & Leather, C. (2000). Modality effects and the development of the word length effect in children. *Memory*, *8*, 1–17. doi:10.1080/096582100387678
- Hitch, G. J., Halliday, M. S., Dodd, A., & Littler, J. E. (1989). Development of rehearsal in short-term memory: Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, *7*, 347–362. doi:10.1111/j.2044-835X.1989.tb00811.x
- Howard, D., & Franklin, S. (1990). Memory without rehearsal. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 287–318). New York: Cambridge University Press. doi:10.1017/CBO9780511665547.016
- Jacquemot, C., Dupoux, E., & Bachoud-Lévi, A.-C. (2011). Is the word-length effect linked to subvocal rehearsal? *Cortex*, *47*, 484–493. doi:10.1016/j.cortex.2010.07.007

- Jalbert, A., Neath, I., Bireta, T. J., & Surprenant, A. M. (2010). When does length cause the word length effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 338–353. doi:10.1037/a0021804
- Jarrold, C. (2013). The development of verbal short-term memory: Insights from, and implications for, adult models. *Paper presented at the Joint Annual Conference of the British Psychological Society Developmental and Cognitive Sections*, Reading, September.
- Jarrold, C., & Citroën, R. (2013). Reevaluating key evidence for the development of rehearsal: Phonological similarity effects in children are subject to proportional scaling artifacts. *Developmental Psychology*, *49*, 837–847. doi:10.1037/a0028771
- Jarrold, C., Danielsson, H., & Wang, X. (2015). Absolute and proportional measures of potential markers of rehearsal, and their implications for accounts of its development. *Frontiers in Psychology*, *6*, 299. doi:10.3389/fpsyg.2015.00299
- Johnson, W., Logie, R. H., & Brockmole, J. R. (2010). Working memory tasks differ in factor structure across age cohorts: Implications for dedifferentiation. *Intelligence*, *38*, 513–528. doi:10.1016/j.intell.2010.06.005
- 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. doi:10.1037/0278-7393.30.3.656
- Lewandowsky, S., & Oberauer, K. (2008). The word-length effect provides no evidence for decay in short-term memory. *Psychonomic Bulletin & Review*, *15*, 875–888. doi:10.3758/PBR.15.5.875
- Lian, A., & Karlsen, P. J. (2004). Advantages and disadvantages of phonological similarity in serial recall and serial recognition of non-words. *Memory and Cognition*, *32*, 223–234. doi:10.3758/BF03196854
- Lobley, K. J., Baddeley, A. D., & Gathercole, S. E. (2005). Phonological similarity effects in verbal complex span. *The Quarterly Journal of Experimental Psychology*, *58A*, 1462–1478. doi:10.1080/02724980443000700
- Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*, 312–319. doi:10.3758/BF03197461
- Logie, R. H. (2011). The functional organisation and the capacity limits of working memory. *Current Directions in Psychological Science*, *20*, 240–245.
- Logie, R. H. (in press). Retiring the central executive. *The Quarterly Journal of Experimental Psychology*. doi:10.1080/17470218.2015.1136657
- Logie, R. H., Della Sala, S., Laiacona, M., Chambers, P., & Wynn, V. (1996). Group aggregates and individual reliability: The case of verbal short-term memory. *Memory and Cognition*, *24*, 305–321. doi:10.3758/BF03213295
- Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *The Quarterly Journal of Experimental Psychology*, *53A*, 626–646. doi:10.1080/713755916
- Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2015). Recalling visual serial order for verbal sequences. *Memory and Cognition*. doi:10.3758/s13421-015-0580-9
- Mora, G., & Camos, V. (2015). Dissociating rehearsal and refreshing in the maintenance of verbal information in 8-year-old children. *Frontiers in Psychology*, *6*, 11. doi:10.3389/fpsyg.2015.00011
- Murray, D. J. (1966). Vocalization-at-presentation and immediate recall, with varying recall methods. *The Quarterly Journal of Experimental Psychology*, *18*, 9–18. doi:10.1080/14640746608400002
- Papagno, C., Lucchelli, F., & Vallar, G. (2008). Phonological recoding, visual short-term store and the effect of unattended speech: Evidence from a case of slowly progressive anarthria. *Cortex*, *44*, 312–324. doi:10.1016/j.cortex.2006.10.001
- Romani, C., McAlpine, S., Olson, A., Tsouknida, E., & Martin, R. (2005). Length, lexicality, and articulatory suppression in immediate recall: Evidence against the articulatory loop. *Journal of Memory and Language*, *52*, 398–415. doi:10.1016/j.jml.2005.01.005
- Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with Kanji materials. *Journal of Memory and Language*, *59*, 1–17. doi:10.1016/j.jml.2008.01.004
- Silveri, M. C., & Baldonero, E. (2013). Visual short-term store can compensate for a defective phonological short-term store in patients with apraxia of speech. *Neuropsychology*, *27*, 628–641. doi:10.1037/a0034033
- Tam, H., Jarrold, C., Baddeley, A. D., & Sabatos-DeVito, M. (2010). The development of memory maintenance: Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology*, *107*, 306–324. doi:10.1016/j.jecp.2010.05.006
- Trojano, L., & Grossi, D. (1995). Phonological and lexical coding in verbal short-term memory and learning. *Brain and Language*, *51*, 336–354. doi:10.1006/brln.1995.1064
- Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a short-term store. *Journal of Verbal Learning and Verbal Behavior*, *23*, 151–161. doi:10.1016/S0022-5371(84)90104-X
- Vallar, A., Basso, G., & Bottini, G. (1990). Phonological processing and sentence comprehension: A neuropsychological case study. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 448–476). New York: Cambridge University Press. doi:10.1017/CBO9780511665547.022
- Vallar, G., Corno, M., & Basso, A. (1992). Auditory and visual verbal short-term memory in aphasia. *Cortex*, *28*, 383–389. doi:10.1016/S0010-9452(13)80148-7
- Vallar, G., Di Betta, A. M., & Silveri, M. C. (1997). The phonological short-term store rehearsal system: Patterns of impairment and neural correlates. *Neuropsychologia*, *35*, 795–812. doi:10.1016/S0028-3932(96)00127-3
- Vallar, G., & Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In A. D. Baddeley, M. D. Kopelman, & B. A. Wilson (Eds.), *Handbook of memory disorders* (2nd ed., pp. 249–270). Chichester: John Wiley & Sons.
- Vallat-Azouvi, C., Weber, T., Legrand, L., & Azouvi, P. (2007). Working memory after severe traumatic brain injury. *Journal of the International Neuropsychological Society*, *13*, 770–780. doi:10.1017/S1355617707070993
- Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2011). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*, 145–160. doi:10.3758/s13421-011-0158-0
- Warrington, E. K., & Shallice, T. (1972). Neuropsychological evidence of visual storage in short-term memory tasks. *Quarterly Journal of Experimental Psychology*, *24*, 30–40. doi:10.1080/14640747208400265
- Waters, G., Rochon, E., & Caplan, D. (1992). The role of high-level speech planning in rehearsal: Evidence from patients with apraxia of speech. *Journal of Memory and Language*, *31*, 54–73. doi:10.1016/0749-596X(92)90005-1