

Serial Recall of Poor Readers in Two Presentation Modalities: Combined Effects of Phonological Similarity and Word Length

Rosemarie S. Irausquin

Tilburg University, The Netherlands

and

Beatrice de Gelder

Tilburg University, The Netherlands, and Université Libre de Bruxelles, Belgium

Immediate ordered memory for words in poor readers was compared with that of two control groups of normal readers, matched on chronological age and reading age, respectively. The groups were equated for basal memory capacity. Phonological similarity and word length were simultaneously manipulated. Items were presented either auditorily (spoken words) or visually (their corresponding drawings). The results suggest that when having to recall a restricted set of items and when verbal output is eliminated, phonological coding and rehearsal occur to the same extent in poor and normal readers, with auditory as well as visual presentation. However, irrespective of presentation modality, absolute performance of the poor readers was still worse than that of their chronological age controls. © 1997 Academic Press

Over the past few decades research on reading disability has converged onto some well-established findings concerning the relationship between impaired reading ability and the processing of linguistic material. More specifically, phonological processing abilities have been found to be deficient in poor

The authors thank all the children and staff of the schools for special education “Hoogvenne” and “De Schans,” and of the primary schools “Gustave School” and “De Sleutel” in Tilburg for their friendly cooperation in this study. We also thank Jean Vroomen for advice in setting up the experiment; Anny Bosman, Ruth Campbell, Monique van Zon, Rebecca Treiman, and two anonymous reviewers for useful comments on earlier drafts of this manuscript; and Antal van den Bosch for making the drawings used in this study.

Address correspondence and reprint requests to Rosemarie Irausquin, Tilburg University, Department of Psychology, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. Fax: +31-13-4662370, E-mail: R.S.Irausquin@kub.nl.

readers. However, the precise nature of the specific deficits and the way in which they relate to reading problems is not yet clear (for a review see Wagner & Torgesen, 1987).

One of the areas in which poor readers often show inferior performance compared to their normal reading peers is verbal short-term memory or working memory (for reviews see Brady, 1991; Jorm, 1983). A measure of short-term memory is memory span, where the task is to recall items in serial order. In the working memory model, as described by Baddeley (1990), a "central executive" acts as a controlling attentional system that supervises and coordinates at least two subsidiary slave systems, the phonological loop and the visuo-spatial sketchpad. The phonological loop is held responsible for the encoding of speech-based information and the visuo-spatial sketchpad for setting up and manipulating visual images. Most relevant with respect to memory for verbal material is the functioning of the phonological loop. This loop comprises two components, a phonological store that holds speech-based information and an articulatory control process engaged in subvocal rehearsal. Memory traces in the phonological store fade away in about $1\frac{1}{2}$ to 2 s, but can be refreshed by subvocal rehearsal, which implies the feedback of information from the store by the articulatory control process. Auditory information gains obligatory access to the phonological store, whereas visual verbal information has to be phonologically recoded by the articulatory control process to gain access to the store.

Several phenomena provide evidence for the loop, among which are the phonological similarity effect and the word length effect (Cowan, Day, Saults, Keller, Johnson, & Flores, 1992; Hulme, 1984; Hulme & Tordoff, 1989; Longoni, Richardson, & Aiello, 1993). The phonological similarity effect refers to poorer recall performance for similar sounding speech units (words or nonwords) than for dissimilar sounding items. Baddeley explains this effect by assuming that in the passive phonological store memory traces for similar sounding words contain fewer discriminable features than those for dissimilar sounding words. The phonemic similarity effect is the clearest source of evidence for the phonological store. The second phenomenon, the word length effect, refers to poorer recall of long words than short ones. Long words take longer to articulate and so fill up more of the limited-capacity loop during rehearsal than do short ones. Consequently fewer long words than short ones can be rehearsed in the same unit of time, leading to relatively more loss of information and thus inferior recall performance for longer words.

This two-process working memory model is not universally accepted and it has been argued that single process models can also account for phonological similarity and word length effects in memory (Hulme & Mackenzie, 1992; Hulme & Tordoff, 1989). Nevertheless, it provides a suitable starting point for exploring immediate memory in relation to reading skill. If dissociated abilities can be found in poor readers for phonemic similarity tasks on the one hand and word length effects on the other,

that might lead us to suspect that different phonological repetition skills contribute to the underlying deficit.

Research on immediate verbal memory has generally shown that the proportion of correctly recalled items is lower in poor readers than in their normal reading peers (Brady, Mann, & Schmidt, 1987; de Gelder & Vroomen, *in press*; de Gelder & Vroomen, 1995; Gathercole & Baddeley, 1990; Pennington, Van Orden, Smith, Green, & Haith, 1990). Furthermore, until recently, it had repeatedly been found that the phonological similarity effect was smaller or even absent in poor readers compared to normal readers (e.g., Mann, Liberman, & Shankweiler, 1980; Shankweiler & Liberman, 1977), irrespective of whether the items to be recalled were presented auditorily or visually. This would indicate either a rather general deficit in the use of phonological coding in working memory or qualitatively poorer phonological representations. However, more recent studies have found that poor readers show similar phonological similarity effects to those of normal readers when presented with list lengths adjusted to memory span (Hall, Wilson, Humphreys, Tinzmann, & Bowyer, 1983; Johnston, Rugg, & Scott, 1987). This suggests that when the difficulty level of the task does not exceed memory capacity, poor readers rely on phonological codes in memory to the same extent as normal readers.

With respect to the lower level of overall performance in poor readers, some studies have sought an explanation in slower articulation speed. A study by Hulme, Thomson, Muir, and Lawrence (1984) provided evidence for the existence of a linear relationship between articulation rate and memory span, found to be constant across different age groups. Articulation rate could thus be a possible candidate for explaining the lower memory span of poor readers. However, other studies have not been supportive in this respect (Brady, Poggie, & Rapala, 1989; Henry & Millar, 1993; Pennington et al., 1990; Stanovich, Nathan, & Zolman, 1988). Brady et al. (1989) found that poor readers were just as fast as normal readers in repeating single words, but not as accurate. This result suggests that it is not so much speed of rehearsal but rather accuracy of reproduction of the presented items that is a more crucial factor. Henry and Millar (1993), in reviewing the evidence on the relationship between rehearsal speed and memory span, state that rehearsal speed is just one of several factors that contribute to verbal memory span. Besides this, they report that much of the correlational evidence linking speech rate to memory span is based on group relationships rather than individual level correlations, which sometimes turn out to be nonsignificant.

A high level of immediate verbal recall is dependent on qualitatively good phonological representations and their use in memory, i.e., phonological coding, as well as adequate rehearsal speed. However, rehearsal cannot take place without phonological coding, and phonological codes need to be refreshed by rehearsal. The aim of this study is to investigate the contribution of these two critical processes of phonological coding and rehearsal in poor readers

compared with chronological age and reading age controls, by a combined manipulation of phonological similarity and word length, respectively.

Two presentation modalities, auditory (spoken words) and visual (drawings of the words), are used to gain further insight into phonological recoding mechanisms. With auditory presentation, phonological coding automatically takes place. The phonological similarity effect, which indicates the use of phonological codes in recall, is therefore expected to occur in this presentation modality. With visual presentation, the effect would only occur when phonological recoding takes place. This presentation modality thus places a heavier load on the ability to use phonological codes in memory than the auditory one, since these codes would have to be generated in the first place. If poor readers do not use phonological codes with visual presentation or use them to a lesser extent than normal readers, the phonological similarity effect should be at least further reduced in poor, but not normal, readers in this presentation condition.

Several factors should be taken into account with respect to certain basic capacities that might directly or indirectly affect the processes under investigation. First, it would be desirable to have an independent measure of rehearsal ability on the one hand and an independent measure of phonological store functioning (irrespective of rehearsal) on the other. Articulation rate is used in this study as a measure of rehearsal ability. This measure also allows us to investigate whether the poor readers have a slower articulation rate than the normal reading controls. Getting a relatively independent indication of phonological store functioning in terms of the duration of memory traces in the store is a more difficult matter. One can never be sure that rehearsal is completely eliminated when trying to get a measure of phonological store functioning. Furthermore, the existence of a separate phonological store is still controversial (Hulme & Mackenzie, 1992; Hulme & Tordoff, 1989). There have been studies that used a nonstrategic memory task or "running memory" task, which comes close to the measure intended here (Cohen & Heath, 1990; Stanovich et al., 1988). Therefore, a similar task is used in this study. This task utilizes the presentation of supraspan lists of digits at a fast presentation rate, to discourage and minimize the use of rehearsal. The participants have to repeat as many digits as possible from the end part of a list in the correct order.

Another factor that should be taken into account refers to basal memory span. As noted before, the studies by Johnston et al. (1987) and Hall et al. (1983) found an equally great phonological similarity effect in poor readers and normal readers when list lengths were adjusted to the memory spans of the respective participants. Also, in a study by Sipe and Engle (1986) on echoic memory processes in good and poor readers, participants were pre-tested to determine their basal memory span (digit span). List length was then adjusted to that span. This was done to avoid floor and ceiling effects and to make the experimental task equally difficult for all participants. Instead of

using different list lengths, in the present study we decided to match the poor readers and their normal reading controls individually on digit span.¹ A further elaboration on this matching procedure is given under Method.

This study presents a detailed investigation of the differences between poor and normal readers in major determinants of phonological memory, through an analysis of the phonological similarity effect, the word length effect, and their interaction. The study is focused on three issues. The first of these, which is also the most central question, is whether the poor readers show smaller phonological similarity effects than their normal reading controls. Based on the results of Johnston et al. (1987) and Hall et al. (1983), one can expect that making the task equally difficult for all matched participants by equalizing basal memory span should lead to similar phonological similarity effects for the poor and normal readers in the auditory as well as in the visual presentation modality. If so, this finding would provide additional support for the suggestion that poor and normal readers use phonological codes to the same extent in short-term memory. A second important issue concerns the effect of word length. If poor readers are deficient in rehearsal abilities, this deficit should be reflected in greater word length effects compared to normal readers, in both the auditory and the visual presentation conditions. Even though earlier studies have advanced rehearsal inefficiency as a possible explanation for lower overall recall in poor readers, the respective studies have been aimed at investigating articulation rate rather than word length effects. The present study takes both of these aspects into account. The third matter of interest concerns an interaction between phonological similarity and word length factors. Results of studies by Hulme (1984) and Hulme and Tordoff (1989) provide support for a relationship between rehearsal rate and the size of the phonological similarity effect. These studies found that the size of the effect increased with age and was correlated with speech rate, even when the effects of age were partialled out. This increase was explained by assuming that rehearsal leads to an additive increase in confusion errors that are fed back to the phonological store. The higher confusability for phonologically similar than dissimilar words consequently leads to a greater difference in error rates between the two kinds of words and this is reflected in a greater absolute difference in recall. If this relationship holds, then it is to be expected that the phonological similarity effect will be smaller for long words than for short ones due to the fact that fewer long words than short ones can be rehearsed per unit of time. Is this relationship the same in poor

¹ A problem with this matching procedure could be that the participants are then made equal with respect to rehearsal abilities. However, referring to the linear relationship found between articulation rate and memory span (Hulme et al., 1984), it could also be the case that this matching procedure would eliminate differences in articulation rate between all three groups of participants, including the younger reading age control group. This turned out not to be the case, as will be described in the results section.

and normal readers? One could argue that, if rehearsal is slower in poor readers, this should be reflected in an even greater reduction of the phonological similarity effect for long words, relative to that in the control groups. So far, there have been no studies on reading disability that compared phonological similarity *and* word length effects in serial recall between poor and normal readers, using auditory as well as visual presentation conditions. Gathercole and Baddeley (1990) investigated both phonological similarity and word length effects in serial recall of language-disordered children who were also poor readers, but these effects were examined separately in different experiments and with only auditory presentation. They found that, except for the longest memory lists, their poor readers were as sensitive to phonological similarity and word length as two control groups that were respectively matched on verbal abilities and nonverbal intelligence. These findings are in accord with those by Johnston et al. (1987) and Hall et al. (1983). The overall performance of the language-disordered children was however lower than that of the control groups, but this could not be attributed to slow articulation rates or impaired perceptual processing.

METHOD

Participants

Three groups of 16 children each participated in the experiment: a group of poor readers and two individually matched control groups, one matched on chronological age (CA control group) and the other on reading age (RA control group). The mean age for the poor readers group was 12 years 3 months (range 11;8–13;0), for the CA control group 12 years 2 months (range 11;9–12;10), and for the RA control group 9 years 5 months (range 8;3–10;5). The poor readers were recruited from primary schools for special education. Participants in the two control groups attended regular primary schools in the same area. Selection was based on performance on the following tests: (a) two reading tests: a word reading test (Eén-minuut-test, Brus & Voeten, 1973) in which isolated words had to be read aloud during one minute, and a pseudoword version of the complementary form of this test, constructed in the laboratory by changing one or more phonemes of each test word (de Gelder & Vroomen, 1991); the reading scores for these two tests represented the number of correctly read words and pseudowords, respectively; (b) forward digit span: subtest of the Wechsler Intelligence Scale for Children—Revised (WISC-R) 1986; (c) verbal IQ: the Peabody Picture Vocabulary Test² (Dunn, 1959; Dutch translation by Manschot & Bonnema,

² For the translated version of this test no standardized norms are available for children older than 8.5 years. To get a reasonable indication of verbal IQ for the Dutch version of the test, 15 points should be subtracted from the American norm IQ score (H. van der Vlugt, personal communication, July 18, 1995). Up to now, there have been no vocabulary tests available in Dutch which are suitable for testing poor readers, i.e., in which words do not have to be read.

1974), a passive vocabulary test in which a child has to indicate the meaning of a word by pointing to the corresponding picture out of four pictures; and (d) nonverbal IQ: Raven's Standard Progressive Matrices (Raven, 1958). Poor readers also received the coding test of the WISC-R before selection, to assess whether they showed any notable deficits in the visual memory and perceptual capacities required for this task. In this task a set of numbered symbols must be copied as accurately and quickly as possible to correspondingly numbered blanks, which are randomly ordered.³ If the standard score on this test was lower than one standard deviation below the standard score for the mean, the child was excluded from the study. Children whose score on the Raven fell below the 25th percentile were also excluded. To make analyses on the Raven scores possible, the percentile scores assigned to the raw scores according to the norms (Raven, Court, & Raven, 1979: table SPM X) were converted into scores from 1 through 8, starting with the lowest (5th) percentile. These converted scores will be referred to as "converted percentile scores" below.

To equate the RA control group and the poor readers group on reading level, mean z scores of the word and pseudoword reading tests were used for individual matching (the correlation between the raw scores, i.e., number of words read correctly, on the two reading tests was high and significant, $r = .87$, $p < .001$). This prevented extreme discrepancies between the matched participants on pseudoword reading ability. For the calculation of mean z scores, the raw scores for, respectively, the word and pseudoword reading tests were first transformed to z scores, for all the children tested before selection ($N = 134$). Matching was based on individual mean z scores, with the restriction that neither the word z scores nor the pseudoword z scores of matched participants differed more than 0.75 (less than the standard deviation of 1 in a z -score distribution).

Participants in all three groups were individually matched on digit span. The criterion of interest for this study was the list length at which a participant started to make mistakes, as this reflects the point at which the task exceeds the memory capacity of the participant. Therefore, instead of the standard criterion to stop testing when a participant fails on both trials of a certain list length, testing was stopped when the participant had made two consecutive errors.⁴ On the basis of this criterion two scores were calculated that reflected: (a) the length of the last list correctly recalled, minus 0.5 points when only

³ According to the manual of the WISC-R, the coding test assesses pattern perception, visual immediate memory, associative learning, ordering, eye-hand coordination, psychomotoric rapidity, attention and concentration, sustained effort, and emotions.

⁴ In the forward digit span test of the WISC-R, lists of increasing length are presented, starting with three items and ending with seven. Of each list length, two trials are presented. Testing standardly stops when a participant makes mistakes on both trials of a certain list length. The number of correctly recalled lists is then transformed into a standard score, as a measure of span performance.

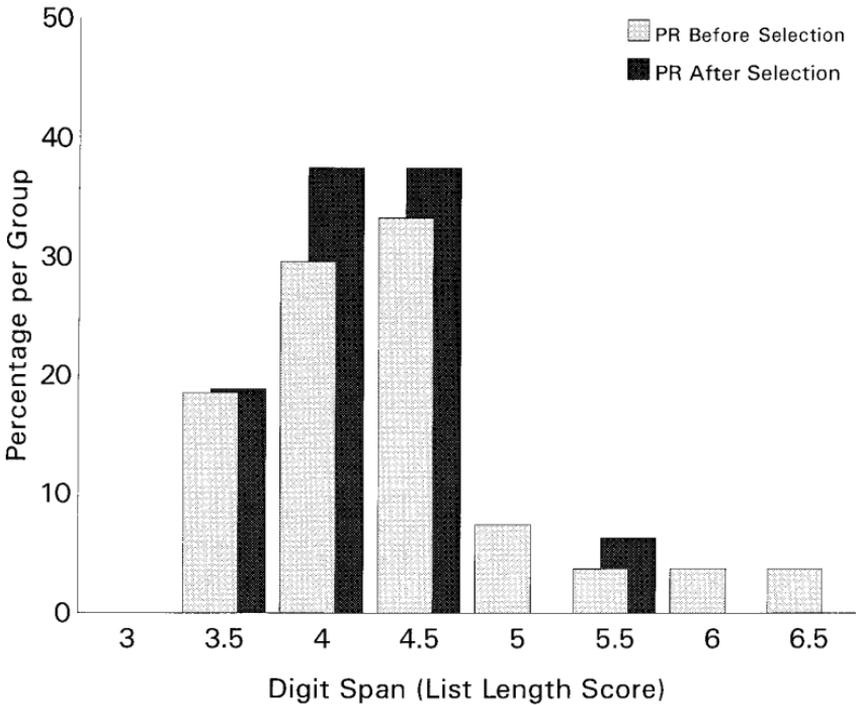


FIG. 1. Frequency distribution of digit span (list length) scores for the poor readers (PR) population sample before selection ($n = 27$) and after selection ($n = 16$).

one of the two lists of that list length was correctly recalled: “list length score”; and (b) the number of correct trials: “correct trials score.” The maximum allowed list length score was 5.50 for each participant, to avoid ceiling effects in the experimental task in which the list length was six. Figures 1 through 3 present the frequency distributions of list length scores before selection (population sample) and after selection, for each group of participants. These figures show that the matching procedure did not lead to the selection of only very poor CA controls or very good RA controls.

Since knowledge of words can have an influence on recall performance for words, the participants were also matched on the verbal IQ score as closely as possible. Table 1 presents the mean scores on the tests mentioned above for each group. One-way analyses with Group (Poor Readers vs CA Control vs RA Control) as the between-subjects variable revealed significant main effects of group on each of the three word reading scores (mean z score, word reading raw score, pseudoword reading raw score: all p values $< .001$). Post hoc comparisons (Newman-Keuls, $\alpha = .05$) showed that the CA control group differed significantly from the poor readers group and the RA control group on each of these variables, whereas these last two groups did not differ from each other. A one-way analysis on the verbal IQ score with Group (PR

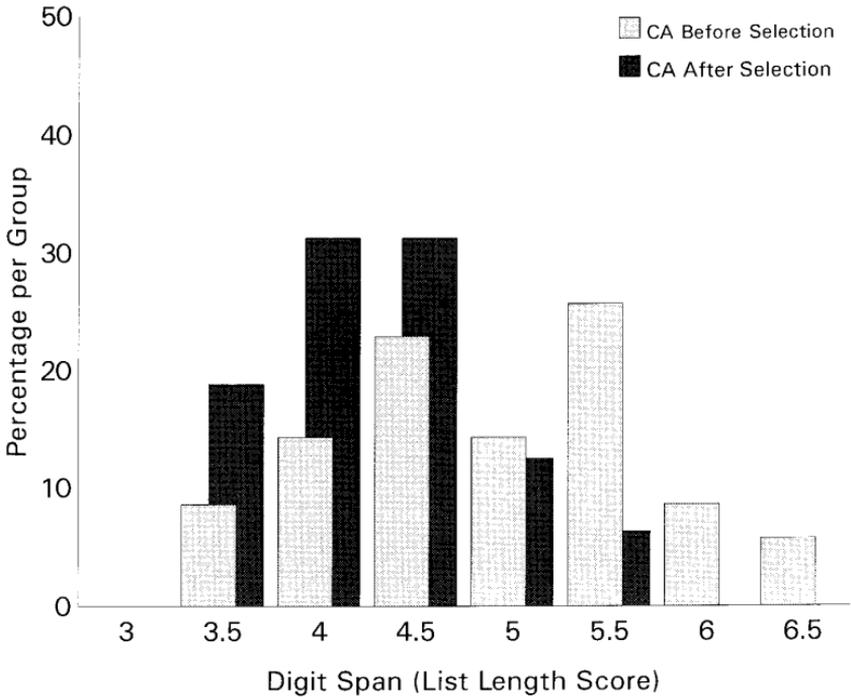


FIG. 2. Frequency distribution of digit span (list length) scores for the CA control group population sample before selection ($n = 35$) and after selection ($n = 16$).

vs CA vs RA) as the between-subjects variable revealed no significant main effect of group.

However, with respect to the converted percentile scores on the Raven, a similar analysis did reveal a significant main effect of group, $F(2,45) = 4.60$, $p < .05$. Post hoc comparisons (Newman-Keuls, $\alpha = .05$) showed that the poor readers group was worse than each of the two control groups, who did not differ significantly from each other on this measure. This finding will be taken into account in the analyses of the experimental data.

Materials

The variables of phonological similarity and word length (one versus three syllables) were orthogonally combined, resulting in the following four word types. Short, nonrhyming words: *zon, das, piek, lip, vel, hut* (meaning “sun, tie, peak, lip, sheet, hut,” respectively); short, rhyming words: *dak, pak, lak, hak, vak, zak* (roof, suit, enamel, heel, partition, sack); long, nonrhyming words: *journalist, elastiek, bioscoop, kapitein, omelet, uniform* (journalist, elastic, cinema, captain, omelet, uniform) and long, rhyming words: *restaurant, muzikant, ledikant, diamant, fabrikant, commandant* (restaurant, musician, bedstead, diamond, factory-owner, commander). All words were mono-

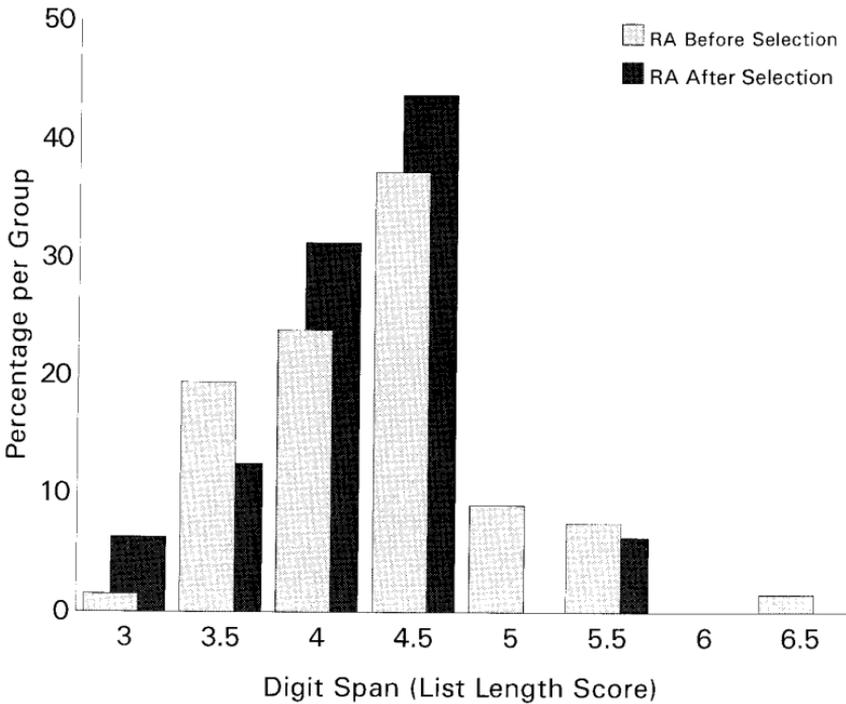


FIG. 3. Frequency distribution of digit span (list length) scores for the RA control group population sample before selection ($n = 67$) and after selection ($n = 16$).

morphemic nouns and the long words had stress on the last syllable. Word frequencies were assessed on the basis of two types of data, (a) the CELEX word frequency lexicon (based on 42 million written words; Baayen, Piepenbrock, & van Rijn, 1993),⁵ and (b) familiarity ratings (on a 7-point scale) by 15 teachers from the four different schools.⁶ The CELEX frequencies did not differ for the four word types, whereas the ratings by the teachers showed that the short words were more familiar than the long words (Newman-Keuls, $\alpha = .05$), irrespective of phonological similarity.

Four types of lists were thus constructed for blocked presentation. For each

⁵ Mean, standard deviation, and range of the frequencies for each of the four word type conditions, as assessed with the CELEX frequency lexicon were as follows. For short, nonrhyming words: M , 763.50; SD , 681.98; range, 107–1808; short rhyming words: M , 1403.67; SD , 1015.74; range, 178–2776; long, nonrhyming words: M , 1113.67; SD , 1493.34; range, 75–4052; long, rhyming words: M , 550.67; SD , 593.54; range, 99–1482.

⁶ Mean, standard deviation and range of the scale-scores (on a 7-point scale) for each of the four word type conditions, as assessed by the ratings of 15 teachers from the different schools were as follows. For short, nonrhyming words: M , 5.04; SD , .91; range, 3.67–7.00; short, rhyming words: M , 5.31; SD , .95; range, 4.00–7.00; long, nonrhyming words: M , 3.67; SD , 1.15; range, 2.00–6.17; long, rhyming words: M , 3.06; SD , .95; range, 4.00–7.00.

TABLE 1
Means for Reading Raw Scores, Digit Span, Verbal IQ,
and Raven Converted Percentile Scores

Group	Reading		Digit span		Verbal IQ	Raven
	Words	Pseudowords	List length	Correct trials		
Poor readers ($n = 16$)						
<i>M</i>	48.38	22.25	4.19	4.31	94.56	4.63
<i>SD</i>	7.53	4.63	0.51	1.08	16.79	0.72
CA control ($n = 16$)						
<i>M</i>	66.88 ^{a,b}	43.88 ^{a,b}	4.28	4.50	105.38	5.63 ^a
<i>SD</i>	6.87	7.78	0.58	1.03	16.17	1.09
RA control ($n = 16$)						
<i>M</i>	45.75	26.00	4.19	4.38	99.81	5.31 ^a
<i>SD</i>	6.13	4.52	0.57	1.15	16.59	1.01

Note. *a*, significantly different ($\alpha = 0.05$) from the respective score in the poor readers group; *b*, significantly different ($\alpha = 0.05$) from the respective score in the RA control group.

block of a certain word type 12 experimental lists were constructed of six items each (all six words of a word type). The items within each list were arranged in quasi-random order, with the restriction that over all 12 lists each item occurred twice in each serial position. Furthermore, 4 practice lists and 2 warm-up lists were constructed with the same items. The 4 practice lists were increasing in length, starting with a list of three items, through six. The warm-up lists were six items in length.

Each word was represented as a black-and-white line drawing, made by a semi-professional artist, on an 8 × 7-cm white card. This resulted in 24 cards to be used as response material, to eliminate verbal output. There were six response cards per word type condition. Six blank cards of the same format were also constructed, which the participant was allowed to use for items that could not be recalled.

For the two presentation modalities, auditory (spoken words) and visual (drawings), stimulus material was constructed in the following way. For the auditory condition, the items were spoken in isolation by a male native speaker of Dutch and recorded on digital audiotape. They were then digitized with a sound editor and played back on digital audiotape. Lists were constructed with an interitem interval of 1 s. Each list was preceded by a 1000-Hz warning signal of 300 ms. After presentation of the last item of each list, there was a pause of 25 s to respond and 10 s for random rearrangement of the response cards. The intertrial interval (ITI) within each block was thus 35 s. A 1200-Hz tone of 200 ms signaled the end of the response time (25 s). For the visual condition, the drawings made for the response cards were scanned for presentation on a computer screen. Lists were constructed in the same way

as for the auditory condition, except that the warning signal generated by the computer before each trial was a 500-Hz tone of 200 ms and the end-of-response time signal was a 1000-Hz tone of 200 ms. Each drawing was presented on screen for 2 s.

Furthermore, for the running memory task, monosyllabic digits from zero through eight (not digit "seven," because it is bisyllabic) were spoken in isolation by the same male speaker of Dutch. Afterward they were digitized and played back on digital audiotape. Lists of these digits were then constructed with a supraspan list length. The lists were 8, 10, 12, 14, or 16 digits in length. A total of 15 experimental lists was constructed, three of each list length. The different list lengths were randomly ordered to prevent predictability of the end of a list. In addition, 3 practice lists were made, one of 12 digits, one of 8 and one of 16. Each digit occurred twice in a list, except for lists of 8 digits (each digit once), to keep the frequency per digit constant. The order of the digits was quasi-random, with the restriction that the repetition of a digit in a list was always separated by at least 2 other digits. The interitem interval within a list was 100 ms and the ITI 10 s. A 1000-Hz warning signal of 300 ms was presented at the beginning of a list and a 1200-Hz tone of 200 ms at the end, which signaled the start of recall.

Procedure

In the first session, participants were assessed on running memory performance and articulation rate for single words. The serial recall experiment took place in the following four sessions, of which the last one also included an articulation rate measurement for pairs of experimental words, at the end of the session. All sessions were on different days. In each serial recall session, two blocks were presented in one of the presentation modalities.

Running Memory Performance

Participants were told that in this test they would always hear a beep first, followed by a series of digits, and at the end another beep. They were instructed not to try to remember all the digits presented, as this would be too difficult, but to try to repeat as many of the last few items from before the end-beep as they could remember, in the order presented. The end-beep was said to be the signal that the list had ended and that repetition could start. Then participants received the three practice trials and the experimenter checked whether they understood the instructions. If necessary, feedback was given. Then the test trials were presented, during which no feedback was given. Responses were written down by the experimenter.

The scoring procedure was as follows. For the last seven serial positions, the number of correct responses per serial position was counted, so that a serial recall curve was obtained for each participant. Furthermore, the slope coefficient over the last three serial positions was calculated using the formula for the calculation of the unstandardized regression coefficient. This slope

coefficient was taken as an indication of the rate of decay of the memory trace in the phonological store, when progressing from last presented items to earlier presented ones. Proportion correct scores per serial position and the slope coefficient thus resulted in two measures on which analyses of variance could be performed.

Articulation Rate

For the assessment of articulatory rate, different words were used for measurement before and after the serial recall experiment. In the first session, three single words of respectively one, two, and three syllables in length were used. These were: "ma," "ballon," and "bakkerij" (meaning "ma, balloon, and bakery," respectively). For the measurement on the last session, four word pairs were used consisting of experimental words used in the serial recall tasks, one pair from each word type. To exclude familiarization with the stimulus material before the experimental task, different and single words were used for the preexperimental measure. In this way, it was also possible to assess whether the two kinds of measurements would show different results.

The word pairs used in this study for the measurement of articulatory rate were selected on the basis of the highest familiarity ratings given by teachers. This was done to prevent influences of unfamiliarity with pronunciation as much as possible, since the words would also not have to be pronounced during the serial recall task. The word pairs were *dak-zak*, *zon-lip*, *restaurant-muzikant*, and *bioscoop-elastiek*. The measurement procedure was the same for both types of stimuli. The experimenter said a word (or word pair) and participants were asked to repeat this word continuously as fast as possible, but with clear pronunciation. For the first word, the experimenter showed the participants how to do the task. They were instructed not to stop repeating until the experimenter said "stop." With each word participants could practice once before the actual measurement was taken. The repetitions were spoken into a microphone connected to a taperecorder. The experimenter made sure that there were at least 10 repetitions of each word or word pair.

After the measurement, the recordings were digitized with a sound editor, which made exact time measurements possible. Two kinds of time measurements were made: (a) the time taken for 10 complete repetitions of the word or word pair, including the time for hesitations (i.e., incomplete repetitions, e.g., *ba. . .bakkerij*), and (b) excluding the time for hesitations. Both measures were converted into number of words per second.

Serial Recall Experiment

Word and response card association training. Before presentation of each experimental block, participants received a training in associating each of the six words used in that block with the corresponding response card out of the six response cards for that block and vice versa. This was done in both the auditory and the visual conditions. Participants were first asked to point to

the correct card belonging to a stimulus word given by the experimenter. This was done with all the six words of a block. They were then asked if there were any words that were unknown to them. If this was the case, the meaning of the respective word was explained to them by the experimenter.⁷ Subsequently, the experimenter would point to each of the response cards in random order and participants were asked to name the word belonging to the respective card that was pointed to. After that, the experimenter would again name each of the six stimulus words in random order and participants were asked to point to the corresponding cards. All participants received at least two naming trials and two pointing trials for each word, but more trials were given if they still made mistakes or took long to name the word.

Auditory condition. The response cards were put in front of the participants, arranged in two columns of three cards in random order. The participants were told that they would hear lists of six words and that for each word there was a response card in front of them. The task was to remember the order in which the words were presented and then put the corresponding cards in the same order. During presentation of the words the response cards were covered with a piece of cardboard. After presentation of the last item of a list the cover was removed and participants had to put the cards in the correct order from left to right. They were not allowed to start with the last presented item or items. Changes in a previously arranged order for the respective test trial were allowed within the response time, but discouraged. Blank response cards could be used for items that could not be recalled, but participants were encouraged to guess.

The participants first received the four practice trials, followed by the two warm-up trials and the experimental trials. Feedback was not given, except for the first practice trial of three items if participants made a mistake there. This was done to make sure that the participants understood the instructions. Responses were written down by the experimenter on a prepared response sheet.

Visual condition. In this presentation modality the same drawings as those on the response cards were now used for presentation on a computer screen. The spoken forms of the words were thus not provided to the participants. Response material and procedure were the same as in the auditory condition. The experimenter made sure that participants were looking at the monitor during presentation of the drawings. They were not allowed to say the stimulus words aloud during presentation.

Design

Half of the participants in each group received the auditory condition first, the other half the visual condition. Each participant who received the auditory

⁷ The words *fabrikant*, *commandant*, *ledikant*, *omelet*, *uniform*, *lak*, and *vak* were unknown to respectively 15, 21, 17, 25, 2, 2, and 2% of all 48 participating children. The higher percentages (over 2%) were grossly distributed over PR:CA:RA in a 2:1:3 ratio. For the low percentages

condition first in a group got a different order of the four blocks of lists (one block for each word type), following a latin square design. Their counterpart in the other half of the group received the same order. For each participant the order of the blocks was identical in both presentation modalities. Matched participants received the same order across conditions.

RESULTS

The results of the running memory task and the articulation rate measures are reported first and their independence is tested by correlational analyses. Next, the results of the serial recall task are presented and related to the running memory and articulation rate measures.

Within-group correlational analyses between the converted percentile scores on the Raven test and mean serial recall performance in each of the four word type conditions per presentation modality showed that only in the CA control group two significant correlations existed. These concerned mean recall performance in the auditory condition for short nonrhyming words ($r = 0.55, p < .05$) and for long nonrhyming words ($r = 0.52, p < .05$) in the same presentation modality. All other correlations were nonsignificant. In a similar correlational analysis, over all three groups, none of the correlations was significant. Therefore, the Raven converted percentile score is not taken up as a covariate in the analyses reported below.

Running Memory Task

Figure 4 shows the mean proportion correct scores for the last seven serial positions in the running memory task, for each group. A 3×7 (Group \times Serial Position) ANOVA on proportion correct scores revealed a marginally significant effect of group, $F(2,45) = 2.45, .09 < p < .1$, and a significant main effect of serial position, $F(6,270) = 529.17, p < .001, \epsilon = 0.4462$.⁸ The interaction was not significant. The marginally significant effect of group was probably due to the lower mean proportion correct scores (over all serial positions) of the poor readers group and the RA control group compared to that of the CA control group (mean proportion correct score for PR: 0.31; RA: 0.32; CA: 0.36). A one-way analysis on the slope coefficient over the last three serial positions (PR: $M = .30, SD = .13$; CA: $M = .22, SD = .11$; RA: $M = .25, SD = .13$), with group as the between-subjects variable, showed that there was no effect of group ($p = .13$). However, the power for this latter test turned out to be too low (just above 0.30). To increase the power to an acceptable level, pairwise group comparisons were made through t tests.

(2%) it was the case that one child in either the poor readers group or the RA control group did not know the word.

⁸ An epsilon adjustment was made to the degrees of freedom for the F -test, using the Greenhouse-Geisser (or Box's) adjustment.

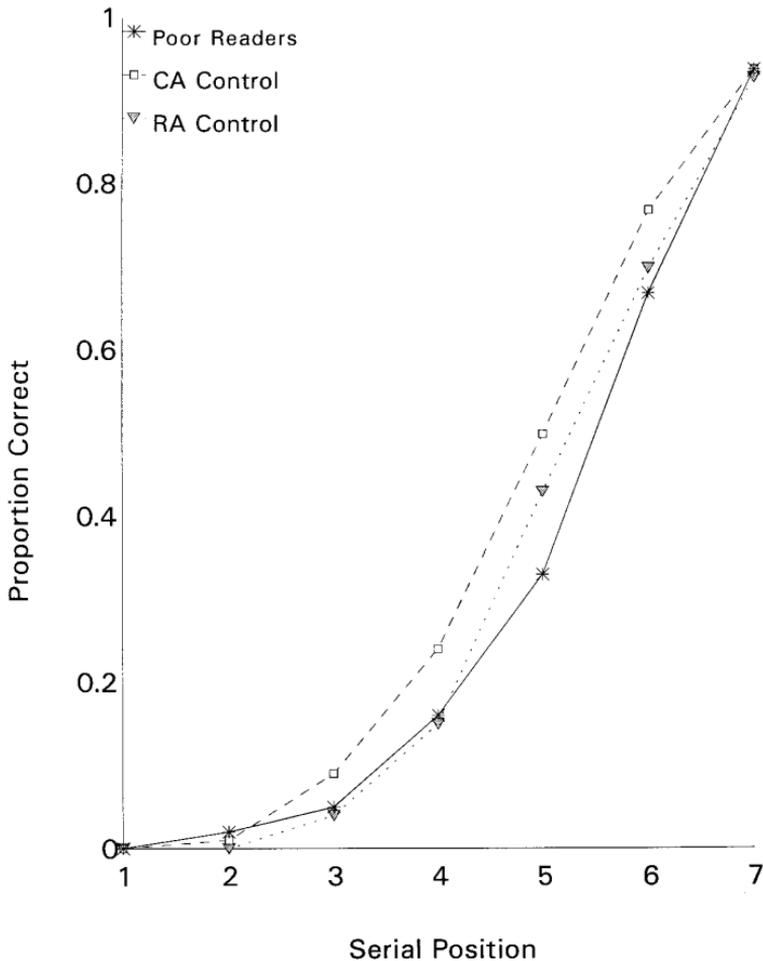


FIG. 4. Running memory performance: mean proportion correct scores for the last seven serial positions. Serial position number 7 represents the last presented digit.

These revealed a significant difference in slope coefficient only between the poor readers group and the CA control group, $t(30) = 2.14$, $p < .05$, implying a steeper slope for the poor readers group.

Articulation Rates

Both measures of articulation rate (i.e., including and excluding hesitations) showed the same results in the between-group analyses for single words as well as for word pairs, except for one interaction in the word pairs analysis, as noted below. However, *for word pairs*, only the measure including hesitations showed a very significant correlation with mean recall performance in the auditory condition, $r = .44$, $p < .01$, as well as in the visual condition, $r = .39$, $p < .01$ (see Table 6). For single words, none of the measures showed

TABLE 2
Mean Articulation Rates (in Words per Second) Including Hesitations, for Single Words

Group	Word			Total
	ma	ballon	bakkerij	
Poor readers				
<i>M</i>	4.84	2.17	1.75	2.92
<i>SD</i>	0.79	0.30	0.15	
CA control				
<i>M</i>	4.62	2.13	1.76	2.84
<i>SD</i>	1.02	0.28	0.15	
RA control				
<i>M</i>	4.23	1.89	1.51	2.55
<i>SD</i>	1.06	0.22	0.22	
Total	4.56	2.06	1.67	2.77

a significant correlation with mean recall performance in either the auditory or the visual condition. Therefore, only the results of the measure including hesitations are reported here.

Single words. Table 2 shows the mean values per single word and per group for the measure including hesitations. A 3×3 (Group \times Word Length) ANOVA revealed significant main effects of group, $F(2,45) = 4.00$, $p < .05$, and word length, $F(2,90) = 403.57$, $p < .001$, $\epsilon = 0.5562$. The interaction was not significant. The main effect of group was due to the fact that both the poor readers group and the CA control group, who did not differ in articulation rate, had significantly higher articulation rates than the RA control group (Newman-Keuls, $\alpha = .05$). The main effect of word length logically reflected slower articulation rates for longer words.

Experimental word pairs. Table 3 shows the mean values per word pair and per group for the measure including hesitations. A $3 \times 2 \times 2$ (Group \times Phonological Similarity \times Word Length) ANOVA revealed significant main effects of group, $F(2,45) = 7.13$, $p < .01$, and word length, $F(1,45) = 1453.81$, $p < .001$ ($M_{\text{short}} = 2.76$; $M_{\text{long}} = 1.31$). The main effect of phonological similarity was marginally significant, $F(1,45) = 2.99$, $p = .09$ ($M_{\text{rhyming}} = 2.07$; $M_{\text{non-rhyming}} = 2$). Significant interactions were found for group by word length, $F(2,45) = 4.25$, $p < .05$, and phonological similarity by word length, $F(1,45) = 4.27$, $p < .05$. Other interactions were not significant. (The group by word length interaction was not significant in the analysis without hesitations.) Again, the main effect of group reflected similar articulation rates for the poor readers group and the CA control group, whereas each of these groups had a significantly faster articulation rate than the RA control group (Newman-Keuls, $\alpha = .05$). The group by word length interaction is depicted in Fig. 5, showing that the increase in articulation rate for short

TABLE 3
 Mean Articulation Rates (in Words per Second) Including Hesitations,
 for the Experimental Word Pairs

Group	Word pair				Total
	Zon-lip	Dak-zak	Bioscoop-elastiek	Restaurant-muzikant	
Poor readers					
<i>M</i>	2.71	2.79	1.38	1.39	2.07
<i>SD</i>	0.40	0.44	0.19	0.22	
CA control					
<i>M</i>	2.86	3.09	1.36	1.38	2.17
<i>SD</i>	0.37	0.50	0.24	0.27	
RA control					
<i>M</i>	2.50	2.61	1.18	1.17	1.86
<i>SD</i>	0.37	0.40	0.18	0.14	
Total	2.69	2.83	1.31	1.31	2.03

words compared to long ones is relatively greater in the CA control group than in the other two groups. The interaction between phonological similarity and word length implied equal articulation rates for the long word pairs, whereas for the short word pairs the rhyming ones were articulated a bit faster than the nonrhyming ones.

Independence of Running Memory Performance and Rehearsal Speed

To test the assumption that performance on the running memory task is unrelated to rehearsal speed, correlational analyses were performed between the running memory performance measures and the articulation rate measures (over all three groups of participants). These analyses may not be exclusive, but should give a reasonable indication whether the assumption above is satisfied.

For purposes of convenience the number of articulation rate values was reduced by calculating mean values for single words and word pairs, respectively, resulting in one mean articulation rate value representing single words and one representing word pairs. Only the results with respect to the articulation rate measure including hesitations will be reported here, because the other measure gave the same results (except in one instance, as will be mentioned below). Prior to calculating these means it was assessed whether the values on which the means were based were themselves significantly and positively correlated. This was the case (all r values were significant with $p < .01$, except between the values for "ma" and "bakkerij": $p < .05$).

Next, a correlational analysis was performed between each of these means and the slope value in the running memory task. None of these correlations was significant. A correlational analysis was also performed between the articulation rate mean measures and the proportion correct scores for the last

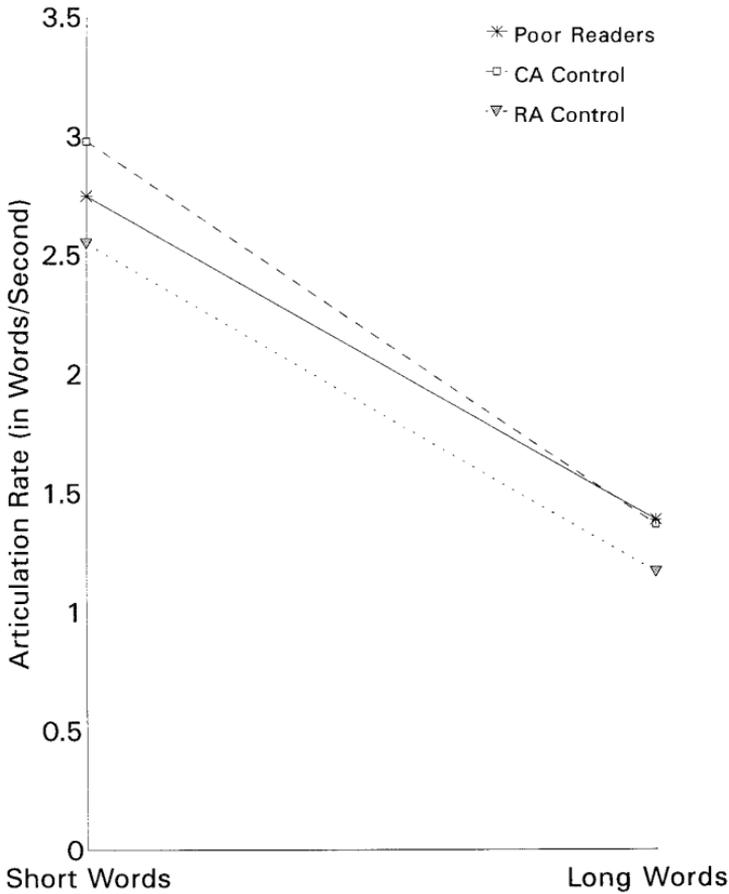


FIG. 5. Mean articulation rates (including hesitations) for short and long word-pairs.

seven serial positions in the running memory task. Only one correlation was significant. That was between the mean articulation rate value for word pairs and the proportion correct score for serial position six (the last but one presented digit): $r = .34, p < .05$). Here the other articulation rate measure, which excluded hesitations, did not show any significant correlation. These results indicate that performance on the running memory task is only marginally related to rehearsal speed.

Serial Recall Experiment

For the auditory condition as well as for the visual condition, proportion correct scores per serial position were calculated first, for each word type block. Consequently, mean proportion correct scores were calculated per word type condition (across serial positions). These scores were used in the analyses reported below.

TABLE 4
Auditory Condition: Mean Proportion Correct Scores per Word Type

Group	Word type				Total
	Short		Long		
	Nonrhyming	Rhyming	Nonrhyming	Rhyming	
Poor readers					
<i>M</i>	.59	.47	.51	.41	.49
<i>SD</i>	.17	.16	.17	.12	
CA control					
<i>M</i>	.66	.54	.57	.55	.58
<i>SD</i>	.20	.15	.18	.15	
RA control					
<i>M</i>	.45	.33	.37	.34	.37
<i>SD</i>	.18	.14	.09	.10	
Total	.57	.45	.48	.43	.48

Tables 4 and 5 present the mean proportion correct scores per word type for the auditory and visual condition, respectively. A $3 \times 2 \times 2 \times 2$ (Group \times Modality \times Phonological Similarity \times Word Length) ANOVA on mean proportion correct scores, with modality, phonological similarity, and word length as within-subject factors, showed that all main effects were significant: group, $F(2,45) = 12.60$, $p < .001$; modality, $F(1,45) = 40.53$, $p < .001$;

TABLE 5
Visual Condition: Mean Proportion Correct Scores per Word Type

Group	Word type				Total
	Short		Long		
	Nonrhyming	Rhyming	Nonrhyming	Rhyming	
Poor readers					
<i>M</i>	.62	.55	.57	.53	.57
<i>SD</i>	.16	.17	.19	.19	
CA control					
<i>M</i>	.77	.63	.69	.60	.67
<i>SD</i>	.13	.17	.18	.22	
RA control					
<i>M</i>	.58	.43	.41	.42	.46
<i>SD</i>	.22	.16	.16	.14	
Total	.66	.54	.56	.52	.57

phonological similarity, $F(1,45) = 28.73, p < .001$; word length, $F(1,45) = 11.96, p < .01$. Only the phonological similarity by word length interaction was significant, $F(1,45) = 11.62, p < .01$.

The main effect of group was attributable to the fact that the RA control group performed worse than the poor readers group, who in turn performed worse than the CA control group (Newman-Keuls, $\alpha = .05$). The main effect of modality reflected better performance in the visual than in the auditory condition. The effect of phonological similarity was due to better recall performance for nonrhyming words than for rhyming words, and the word length effect implied better performance for short words as opposed to long ones. The interaction between phonological similarity and word length was further explored using simple effects analyses, which showed that the interaction was mainly due to a nonsignificant word length effect for rhyming words and a significant word length effect for nonrhyming words, $F(1,45) = 19.98, p < .001$. Conversely, the effect of phonological similarity was present for short words as well as for long ones, $F(1,45) = 37.00, p < .001$ and $F(1,45) = 6.34, p < .05$, respectively, but tended to be greater for short words.

Relations between Serial Recall Performance and Running Memory and Articulation Rate

To assess how running memory performance and articulation rate related to serial recall performance, correlational analyses were performed between mean recall performance in the auditory and visual conditions of the serial recall task and the running memory and word pair articulation rate measures. Table 6 shows the results of these analyses per group and across all groups. The correlations with respect to articulation rate for single words are not shown, since none of these correlations was significant. The analyses show that within the poor readers group and the RA control group, serial recall performance is significantly related to recall performance in the running memory task. It is also apparent that in the poor readers group the correlation between mean serial recall performance in each presentation condition and the slope coefficient for the running memory task is highest for all groups, though not significant. Furthermore, in the overall analysis across groups, the articulation rate measure including hesitations shows a much stronger relationship with mean serial recall performance than the measure excluding hesitations. The latter only showed a relation with recall performance in the auditory condition. Articulation rate for word pair repetitions including hesitations thus seems to be the most representative measure for rehearsal speed in relation to serial recall performance. Remarkable is the fact that in the CA control group almost all correlations are very low.

DISCUSSION

The present study compared the serial recall performance of a group of poor readers to that of their reading age and chronological age-matched controls in

TABLE 6

Correlations between Mean Proportion Correct Scores for the Auditory and Visual Conditions and Running Memory and Word Pair Articulation Rate Mean Scores

	Running memory proportion correct	Running memory slope coefficient	Articulation rate for word pairs, including hesitations	Articulation rate for word pairs, excluding hesitations
Poor readers ($n = 16$)				
Auditory	.56*	-.34	.44	.33
Visual	.55*	-.32	.25	.08
CA control group ($n = 16$)				
Auditory	.16	-.02	.02	-.07
Visual	.09	.07	-.12	-.23
RA control group ($n = 16$)				
Auditory	.63**	-.06	.25	.12
Visual	.61*	-.13	.38	.28
All groups ($n = 48$)				
Auditory	.47**	-.16	.44**	.32*
Visual	.48**	-.17	.39**	.27

* $p < .05$, two-tailed; ** $p < .01$, two-tailed.

an auditory and a visual presentation modality. All groups were carefully screened for equivalent digit span to control for basal memory capacity. Independent measures of speech rate and phonological running memory were taken to evaluate the contribution of respectively rehearsal speed and (presumably) phonological store functioning. The aim was to investigate whether, when controlling for memory capacity, differences in performance between the poor readers and their controls could be attributed mainly to deficits in one or both of the processes of phonological coding and subvocal rehearsal, or in their interaction. For this purpose, items were constructed in such a way that phonological similarity and word length were simultaneously manipulated. The visual presentation condition was used to assess whether poor readers use phonological codes in a condition where these codes are not provided and, if so, whether they use them to the same extent as normal readers. The main findings of the serial recall task will be discussed first, followed by discussion of the running memory and speech rate results. Finally, other interesting but less relevant findings will be addressed.

The results of the serial recall task suggest that when presented with a

restricted set of items to recall and when basal memory span is taken into account, phonological similarity and word length effects occur to the same extent in poor and normal readers and also show a similar interaction. This is the case with auditory as well as with visual presentation of the items. We can thus conclude that under the controlled conditions mentioned above, poor readers seem to use phonological codes and rehearse to the same extent as normal readers in a serial recall task, even when phonology is not provided. These findings do not support the original claim made by Mann et al. (1980) and Shankweiler and Liberman (1977) that poor readers show a reduced phonological similarity effect. The results do support the more recent studies by Johnston et al. (1987) and Hall et al. (1983) in which comparable phonological similarity effects were found for good and poor readers, even with visual presentation of the recall items. In addition, also with respect to word length effects, the findings are in accord with those of Gathercole and Baddeley (1990) where both phonological similarity and word length effects were comparable across poor and normal readers, at least for list lengths that were not too long.

Another important finding, which was unexpected, is that the poor readers' overall recall performance for the words (irrespective of presentation modality) was lower than that of the CA control group, even though their digit span was the same. At present, there is no clear explanation for this finding. The majority of studies on serial recall performance in poor readers—even those manipulating task or subject factors in such a way that recall strategies, rehearsal abilities, or extent of phonological coding were similar for poor and normal readers—found lower overall performance for poor readers compared to normal readers (for reviews see Brady, 1991, and Jorm, 1983; de Gelder & Vroomen, in press; Hall et al., 1983; Watson & Engle, 1982). In the study by de Gelder and Vroomen (in press), adult dyslexic subjects did not differ from normal reading adults in effects of presentation modality (written words, spoken words, sounds or pictures of the same stimuli), but did show lower overall performance. In the present study, the fact that the poor readers performed worse on word recall than the CA control group despite similar digit spans suggests a contribution from long-term memory as the most plausible explanation. Words are less familiar and usually phonologically more complex than the highly overlearned and practiced digits (from one to nine). Slower activation or less accurate representations of phonological codes in long-term memory, which has often been found in poor readers (Ackerman, Dykman, & Gardner, 1990; Bowers & Swanson, 1991; Brady et al., 1989; Wolf, 1991), may play a stronger role with words than with the more familiar digits (Hulme et al., 1991). Phonological codes have to be accessed in the recognition of words from speech input, as in the auditory condition in this study, and when engaging in phonological recoding of visual stimuli, as in the visual condition and when reading. Furthermore, during the rehearsal process in serial recall, with either auditory or visual presentation, phonological codes have to be

refreshed in the phonological store by the articulatory rehearsal process. Long-term memory for the phonological forms of words may make an important contribution to the reconstruction of (partially decayed) information held in the loop (Hulme et al., 1991).

Other explanations for the unexpected result may concern phonological store functioning or nonverbal IQ. However, these explanations seem much less plausible. Running memory, in terms of decay rate of the memory trace in the phonological store, did show a relation to serial recall performance, but only weak group differences were found on running memory performance. Nonverbal IQ was lower for the poor readers than for the CA control group, and within the latter group this measure correlated with recall performance in two of the auditory conditions. This suggests that the superior recall performance of the CA control group was in part due to better (nonphonological) recall strategies. However, since the groups did not differ in phonological similarity or word length effects and there was no overall correlation between nonverbal IQ and recall performance, this factor does not seem to play an important role.

The results of the running memory and articulation rate tasks showed the following. Performance on the running memory task provides some support for faster decay of the memory trace in the phonological store for poor readers compared to their chronological age controls, although the group difference was not very strong. For the younger reading age controls performance on this task was intermediate between that of the poor readers and the CA control group. Running memory was not importantly related to articulation rate. The study by Cohen and Heath (1990) is supportive in this respect. Thus, running memory performance seems more related to phonological store functioning than to rehearsal abilities.

The poor readers group and the CA control group did not differ in articulation rate. This is in accord with other studies finding no support for slower articulation rates in poor readers, as mentioned in the introduction. However, the younger RA control group exhibited a slower speech rate for words than the two older groups, even though all groups had been matched on digit span. This result is probably due to the overlearned nature of digits, making them better retrievable than words (Henry & Millar, 1993; Hulme, Maughan, & Brown, 1991).

Less relevant with respect to the present study, but interesting to note, are the following findings. Recall was better in the visual condition than in the auditory one, for all groups. This may have resulted from the longer presentation times in the visual condition, leaving more time for rehearsal. With respect to overall recall, the CA control group performed better than the poor readers group, who in turn performed better than the RA control group. The fact that the RA control group had the lowest overall performance in both presentation conditions, even though they had been matched on digit span, can be explained by their slower articulation rate for the words used. Also,

familiarity with words normally increases as one gets older. This may also have contributed to the better performance of the older groups.

A remarkable finding was the fact that, despite a slower articulation rate, the RA control group did not show greater word length effects compared to the other two groups who did not differ in speech rate. This result raises the question whether speech rate really is a good measure for rehearsal speed, if it is assumed that rehearsal speed is related to the size of the word length effect. There is indeed some controversy with respect to this assumption, in the sense that articulatory rehearsal itself may not be the only factor contributing to word length effects. Caplan, Rochon, and Waters (1992) state that activation of phonological forms alone, for phonological output planning, appears to be sufficient to produce the word length effect.⁹ Another factor that may have obscured a possible relationship between articulation rate and word length effects is the fact that the long words were rated less familiar for the children than the short words.

A second unexpected finding with respect to articulation rate was that the measure including hesitations showed a much stronger relation to serial recall performance than the measure without hesitations. A possible explanation for this result may be the following (see also Levelt, 1989). Hesitations occur when covert or overt errors are detected during the pronunciation of the words. These errors may result from confusions between activated phonological segments during the phonological encoding process, i.e., when the phonological forms of the words are planned for motoric output. Since the articulation rate task requires fast repetition of always the same items, the phonological segments of these items will be activated and deactivated in rapid succession. Failures may occur in this activation–deactivation process, leading to misselections of activated units at the segmental or phonetic spellout level. This process of phonological word form planning also takes place during rehearsal in a serial recall task. Better functioning of the control processes regulating activation and deactivation of sublexical units during phonological encoding should lead to less confusion during rehearsal and thus also better memory span performance. The articulation rate measure including hesitations may therefore be a more representative measure of rehearsal efficiency.

Even though the performance of the groups did not differ in terms of the interaction between the effects of phonological similarity and word length, it is interesting to note that indeed an interaction existed between these two factors. The interaction conformed to the assumption that when rehearsal is slower, as in the case of rehearsal of long words compared to short ones, the

⁹ In the present study, no relationship was found within any of the three groups of participants, between the size of the word length effects (for nonrhyming words and rhyming words separately) and any of the mean articulation rate measures for single words or word pairs, neither in the auditory nor the visual condition.

phonological similarity effect is smaller (Hulme, 1984; Hulme & Tordoff, 1989). However, this explanation does not seem applicable here, since the RA control group exhibited slower articulation rates but phonological similarity effects similar to those of the faster speaking older groups. Another explanation with respect to the present study is that the proportion of phonological overlap was smaller among the long rhyming words than among the short rhyming words. Furthermore, the interaction also implied absence of a word length effect for rhyming words, whereas the nonrhyming words did show a word length effect. The phonological similarity of the rhyming items may have overruled the word length effect for these items.

Before drawing a final conclusion from the results of the present study, attention should be paid to the very strict selection procedure that has been followed and its consequence for the generalizability of the results. In addition to being matched with the control groups on several variables, including verbal IQ and digit span, the poor readers also had to perform at least at an average level on some other variables. These concerned nonverbal IQ and the score on the coding test. Given these stringent criteria, one might of course wonder whether the groups may be considered representative samples of their respective populations. Poor readers' verbal IQ and digit span are usually lower than that of normal readers. Nonverbal IQ has not been shown to be consistently related to reading ability and may range considerably within the population of poor readers. The coding test requires capabilities that poor readers often perform poorly on (see Footnote 3). However, since the study was aimed at investigating short-term memory performance in relation to reading disability, the best way to do this is to exclude factors known to influence either memory performance as such, or reading ability. Controlling for these factors clearly also leads to exclusion of a subpopulation of poor readers with shortcomings other than or in addition to those of the selected group. However, the results of the present study show that even when all these factors are controlled for, while the groups remain differentiated by reading ability, the poor readers group still shows lower overall recall performance than their normal reading peers. This result cannot be explained by poorer phonological coding or rehearsal in the poor readers group. This finding has general implications for understanding phonological memory of poor readers, extending also to groups of poor readers that are defined by less stringent or less transparent criteria than the ones used in this study. As shown here, deficits in the use of phonological codes or rehearsal efficiency do not seem to lie at the core of the lower short-term memory span of poor readers.

To conclude, the findings suggest that when basal memory span is taken into account, poor readers engage in phonological coding and rehearsal to the same extent as normal readers of the same age or reading level, at least when presented with a restricted set of stimuli to recall and when verbal output processes are eliminated. This holds for auditory as well as visual presentation of the items to be recalled. However, the overall performance

of the poor readers remains lower than that of their normal reading age mates. Rehearsal and the use of phonological codes, which have earlier been advanced as possible sources of the lower memory span of poor readers, thus do not appear to have sufficient explanatory strength. Future research might therefore direct attention to other factors influencing overall recall performance in poor readers. Long-term memory for the phonological forms of words is a possible candidate in this respect.

REFERENCES

- Ackerman, P. T., Dykman, R. A., & Gardner, M. Y. (1990). Counting rate, naming rate, phonological sensitivity, and memory span: Major factors in dyslexia. *Journal of Learning Disabilities*, *5*, 325–319.
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove: Erlbaum.
- Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability: Multiple measures of a singular process. *Journal of Experimental Child Psychology*, *51*, 195–219.
- Brady, S. A. (1991). The role of working memory in reading disability. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 129–151). Hillsdale, NJ: Erlbaum.
- Brady, S., Mann, V., & Schmidt, R. (1987). Errors in short-term memory for good and poor readers. *Memory and Cognition*, *15*, 444–453.
- Brady, S., Poggie, E., & Rapala, M. (1989). Speech repetition abilities in children who differ in reading skill. *Language and Speech*, *32*, 109–122.
- Brus, B. Th., & Voeten, M. J. M. (1973). *Eén-minuut-test vorm A en B. Verantwoording en handleiding*. Nijmegen, The Netherlands: Berkhout.
- Caplan, D., Rochon, E., & Waters, G. (1992). Articulatory and phonological determinants of word length effects in span tasks. *The Quarterly Journal of Experimental Psychology A*, *45*, 177–192.
- Cohen, R. L., & Heath, M. (1990). The development of serial short-term memory and the articulatory loop hypothesis. *Intelligence*, *14*, 151–171.
- Cowan, N., Day, L., Saults, J. S., Keller, T. A., Johnson, T., & Flores, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory and Language*, *31*, 1–17.
- de Gelder, B., & Vroomen, J. (1991). Phonological deficits: Beneath the surface of reading-acquisition problems. *Psychological Research*, *53*, 88–97.
- de Gelder, B., & Vroomen, J. (in press). Modality effects in immediate recall of verbal and non-verbal information in adult developmental dyslexics. *European Journal of Cognitive Psychology*.
- de Gelder, B., & Vroomen, J. (1995). Memory deficits for heard and lipread speech in young and adult poor readers. In B. de Gelder & J. Morais (Eds.), *Speech and reading: A comparative approach*. Hove: Erlbaum.
- Dunn, L. M. (1959). *Peabody Picture Vocabulary Test*. Washington: American Guidance Service.
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*, 336–360.
- Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B., & Bowyer, P. M. (1983). Phonemic-similarity effects in good versus poor readers. *Memory and Cognition*, *11*, 520–527.
- Henry, L. A., & Millar, S. (1993). Why does memory span improve with age? A review of the evidence for two current hypotheses. *European Journal of Cognitive Psychology*, *5*, 241–287.

- Hulme, C. (1984). Developmental differences in the effects of acoustic similarity on memory span. *Developmental Psychology*, **20**, 650–652.
- Hulme, C., & Mackenzie, S. (1992). *Working memory and severe learning difficulties*. Hove: Erlbaum.
- Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, **30**, 685–701.
- Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology*, **38**, 241–253.
- Hulme, C., & Tordoff, V. (1989). Working memory development: The effects of speech rate, word length and acoustic similarity on serial recall. *Journal of Experimental Child Psychology*, **41**, 61–75.
- Johnston, R. S., Rugg, M. D., & Scott, T. (1987). Phonological similarity effects, memory span and developmental reading disorders: The nature of the relationship. *British Journal of Psychology*, **78**, 205–211.
- Jorm, A. F. (1983). Specific reading retardation and working memory: A review. *British Journal of Psychology*, **74**, 311–342.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Longoni, A. M., Richardson, J. T. E., & Aiello, A. (1993). Articulatory rehearsal and phonological storage in working memory. *Memory and Cognition*, **21**, 11–22.
- Mann, V. A., Liberman, I. Y., & Shankweiler, D. (1980). Children's memory for sentences and word strings in relation to reading ability. *Memory and Cognition*, **8**, 329–335.
- Manschot, W., & Bonnema, J. Th. (1974). *Handleiding bij de experimentele Nederlandse normering van de Peabody Picture Vocabulary Test*. Amsterdam, The Netherlands: Swets & Zeitlinger.
- Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Development*, **61**, 1753–1778.
- Raven, J. C. (1958). *Standard progressive matrices*. London: Lewis.
- Raven, J. C., Court, J. H., & Raven, J. (1979). *Manual for Raven's progressive matrices and vocabulary scales*. London: Lewis.
- Shankweiler, D., & Liberman, I. Y. (1977). Exploring the relations between reading and speech. In R. M. Knights & D. J. Bakker (Eds.), *The neuropsychology of learning disorders: Theoretical approaches* (pp. 297–313). Baltimore: University Park Press.
- Sipe, S., & Engle, R. W. (1986). Echoic memory processes in good and poor readers. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **12**, 402–412.
- Stanovich, K. E., Nathan, R. G., & Zolman, J. E. (1988). The developmental lag hypothesis in reading: Longitudinal and matched reading-level comparisons. *Child Development*, **59**, 71–86.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, **101**, 192–212.
- Watson, E. S., & Engle, R. W. (1982). Is it lateralization, processing strategies, or both that distinguishes good and poor readers? *Journal of Experimental Child Psychology*, **34**, 1–19.
- Wechsler, D. (1986). *Wechsler Intelligence Scale for Children—Revised, Nederlandstalige uitgave*. Lisse, The Netherlands: Swets & Zeitlinger.
- Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly*, **26**, 123–141.