Applied Psycholinguistics **20** (1999), 563–588
Printed in the United States of America

The development of phonemic coding strategies for serial recall

SUSAN NITTROUER and MARNIE E. MILLER
Boys Town National Research Hospital

ADDRESS FOR CORRESPONDENCE
Susan Nittrouer, Boys Town National Research Hospital, 555 North 30th Street, Omaha, NE 68131. Email: nittrouer@boystown.org

ABSTRACT
This study examined differences between adults and children and between normal and poor readers in the use of phonemic coding strategies for storing words in working memory. In the first experiment, adults, 11-year-olds, and 8-year-olds (categorized as normal or poor readers) recalled eight-item strings of rhyming and nonrhyming words. A developmental decrease in errors was observed for adults, 11-year-olds, and normal-reading 8-year-olds that reflected an improvement in the phonemic coding of items in working memory, but no difference was found between normal- and poor-reading 8-year-olds in the use of phonemic coding strategies. A second experiment with shorter lists and more children supported the latter finding. The results were interpreted as demonstrating that the ability to access syllable-internal phonemic structure is a necessary precursor to the development of phonemic coding strategies for working memory, but that the use of that structure for storing words in working memory is a skill that develops independently and later than the ability to access phonemic structure.

The comprehension of language, whether spoken or written, requires that the receiver retain a string of words in memory long enough to perform appropriate linguistic analyses (most notably, syntactic analysis). Numerous studies have supported the notion that adults' abilities to retain sufficiently long strings of words in the order presented derive from their abilities to process the signal such that phonological information is extracted and then that information is used for storing words in working memory (e.g., Baddeley, 1966; Baddeley, Thomson, & Buchanan, 1975; Campbell & Dodd, 1980; Spoehr & Corin, 1978). Thus, the input and storage of linguistic materials in working memory using a phonological code is considered critical for further linguistic processing. Research on this topic has frequently investigated the serial recall of linguistic items. That is, strings of linguistic items (such as digits, letters, nonsense syllables, or words) are presented either auditorily or visually (as print or pictures), and participants must recall the items in the order in which they were presented. When items with confusible phonological structures (such as rhyming words) are used, more errors are made across all list positions than when items with phonologically nonconfusible structures are used (Baddeley, 1966; Conrad & Hull, 1964; Salame & Baddeley, 1986). Montgomery (1995a) attributed the benefits shown for phonologically nonconfusible items to the fact that they are represented more distinctly in memory. Thus, when linguistic items can be and are entered into working memory using a phonological code, the items themselves, as well as the order of the items, are better retained than when a phonological code is not used or fails to distinguish among items very well.

These findings lead to the prediction that individuals with phonological processing problems (i.e., problems accessing and manipulating phonological structure) would fail to show an advantage in serial recall for phonologically nonconfusible materials over phonologically confusible materials.[1] Although a variety of tasks can be used to examine phonological processing abilities, individuals with reading disabilities demonstrate well-replicated difficulties and/or developmental delays with virtually all such tasks (Stanovich, Cunningham, & Cramer, 1984; Wagner, 1986; Wagner & Torgesen, 1987). As a result, poor readers provide an excellent opportunity for testing the prediction offered above that such individuals would fail to show an advantage in serial recall for phonologically nonconfusible items over phonologically confusible items. In fact, some studies with children classified according to reading ability have provided support for that prediction. Shankweiler, Liberman, Mark, Fowler, and Fischer (1979) compared recall for nonrhyming and rhyming letter strings by three groups of children near the end of second grade: "superior" readers, whose reading abilities were nearly at a fifth grade level; "marginal" readers, whose reading abilities were at a mid-second grade level; and "poor" readers, whose reading abilities were at a beginning second grade level. Letter strings were presented in two different manners: visually and auditorily. Regardless of mode of presentation, the superior readers showed a greater advantage (i.e., fewer errors) for the nonrhyming letters than for the rhyming letters than did either of the other groups. A statistically significant Reading Ability × Rhyme interaction supported this conclusion. Mann and Liberman (1984) demonstrated essentially the same interaction for words presented auditorily to children in kindergarten and first grade with "good," "average," and "poor" word-attack skills. Similarly, Spring and Perry (1983) found an interaction of reading ability and rhyme for pictures of simple nouns presented to children in third, fourth, and fifth grades who were classified as "good" or "poor" readers. However, other studies have failed to show a Reading Ability × Rhyme interaction in serial recall tasks. For example, in a series of five experiments, Hall, Wilson, Humphreys, Tinzmann, and Bowyer (1983) found that "normal" and "poor" readers in second, third, and fourth grades demonstrated the same pattern of errors across rhyming and nonrhyming letters and monosyllabic words, whether presented visually or auditorily. However, in two experiments the overall group effect was significant or approached significance, demonstrating that poor readers generally made more errors in recall. Pennington, Van Orden, Smith, Green, and Haith (1990) investigated serial recall for nonrhyming and rhyming monosyllabic words presented auditorily to two groups of adult dyslexics: those with a family history of dyslexia (familial dyslexics) and those without (clinic dyslexics). Each dyslexic group was matched with its own control group of same-age, normal-reading peers. Compared to their own controls, the clinic dyslexics demonstrated more recall errors overall; the familial dyslexics did not. Of most interest, Pennington et al. reported no hint of a Reading Ability × Rhyme interaction.

In these experiments no indication is given as to the kind of errors that poor readers tended to make: errors of item recall or of order recall. Item errors refer to those errors in which the participant cannot recall what items were presented. With order errors, the participant correctly remembers the items but has the order of presentation wrong. Brady, Shankweiler, and Mann (1983) examined recall of serial order for auditorily presented lists of nonrhyming and rhyming words by third graders who were separated into two reading groups: "good" readers, whose reading abilities were at almost a sixth grade level, and "poor" readers, whose reading abilities were at a late second grade level. The participants were asked to repeat the words in the order in which they were presented. Brady et al. found that the poor readers made more errors overall in recall than the good readers, but there was no Reading Ability × Rhyme interaction. However, when these authors scored the correctness of responses without regard for correctness of order, they found that there was a Reading Ability × Rhyme interaction. This interaction was traced to a higher proportion of transposition errors given by poor readers to the nonrhyming materials (e.g., *trait, plane* instead of *train, plate*). In other words, the poor readers made more errors in recall of items per se than in recall of order, compared to the good readers. This trend suggests that the poor readers had greater difficulty storing items in working memory with a phonological code, but it clouds the interpretation that it is specifically the recall of order that suffers from this difficulty.

The question of the kind of error that poor readers tend to make is important because it could affect our interpretation of the relation between phonological coding for serial recall and sentence comprehension. Individuals with reading disabilities often display difficulties comprehending sentences with complex syntax (e.g., Byrne, 1981; Smith, Mann, & Shankweiler, 1986; Stein, Cairns, & Zurif, 1984; Vogel, 1975). This difficulty has been attributed by some investigators precisely to the poor phonological processing abilities of poor readers rather than to a syntactic deficit. Specifically, the suggestion concerning the problem in verbal working memory of poor readers derives from modular views of language processing. Modular views hold that information (such as a linguistic signal) is processed in a strictly bottom-up manner and is encapsulated so that information from other systems cannot influence processing. If poor readers have difficulty accessing phonological structure in the stimuli, this difficulty will create a "bottleneck" that impedes the flow of information through the module to "higher levels," which presumably include working memory (Bar-Shalom, Crain, & Shankweiler, 1993; Crain, 1989; Crain & Shankweiler, 1991; Crain, Shankweiler, Macaruso, & Bar-Shalom, 1990; Mann, Cowin, & Schoenheimer, 1989; Smith, Macaruso, Shankweiler, & Crain, 1989; but cf. Byrne, 1981; Stein et al., 1984). This interpretation predicts both more item errors (as Brady et al., 1983, observed) and more order errors. However, order errors are of more interest when trying to relate serial recall and sentence comprehension. A greater preponderance of order errors by poor readers needs to be demonstrated if phonological processing deficits are to be posited as the source of observed differences in sentence comprehension between poor and good or normal readers. This requirement follows from the fact that in most tasks testing sentence comprehension, as in real-world communication situations, the items to be stored in working memory are obvious. Common tasks used to investigate sentence comprehension include the object manipulation task (in which participants manipulate small objects according to a sentence) and the sentence–picture matching task (in which participants point to the picture, out of several, that correctly depicts the action in the sentence). Transposition errors of the kind described by Brady et al. (1983) are close to impossible in these tasks because the objects performing the actions or being acted upon are obvious.

Only one study has explicitly explored order errors in serial recall by poor readers. Katz, Shankweiler, and Liberman (1981) used pictures of real items and of "doodle" drawings as stimuli, thus constraining the set of items to be remembered. Participants saw the pictures presented sequentially on a screen; they were asked to recall the order of presentation by arranging cards with the same pictures on them. The participants were second graders categorized as "good" or "poor" readers, based on their word recognition skills. Results showed that children in both groups performed similarly for the doodles (i.e., items that could not be coded phonologically in working memory). Both groups showed an improvement in order recall for the real items (i.e., items that could be coded phonologically in working memory), but good readers showed more improvement than poor readers. Thus, some support has been garnered for the suggestion that poor readers have more difficulty than good readers specifically in recalling the order of phonologically codable items. In turn, this suggestion bolsters the argument that poor readers' difficulties with sentence comprehension can be explained largely by problems with using a phonological code for entering and/or storing items in working memory. However, Katz et al. used stimuli that categorically were or were not codable as phonological sequences, and both groups showed an advantage for the phonologically codable items; it was only the magnitude of the advantage that differed. Thus, it is possible that both groups used a phonological code to the same extent, but that the use of that code provided more benefit for the good readers than for the poor readers. Be that as it may, a study by Montgomery (1995b) provided support for the general suggestion that the sentence comprehension difficulties experienced by poor readers can be traced to problems with using phonological codes for linguistic processing. In that study, children with specific language impairments were found to have more difficulty than their normal-language peers in repeating multisyllabic nonsense words, a task that presumably requires access to phonological structure in order to derive a representation. Furthermore, a statistically significant correlation of .62 was found between scores on this nonsense repetition task and those on a task of sentence comprehension.

The primary purpose of this study was to explore further two factors that may affect the order recall of linguistic materials: reading ability and age. The central hypothesis to be tested was that the use of a phonological code for storing linguistic items in working memory is a skill that emerges automatically with the development of lower level phonological skills (i.e., the ability to recognize phonological structure in linguistic signals). That is, once a child can recognize the phonological structure of the linguistic signal being heard or read, the use of that code to enter linguistic signals into working memory may be automatic. An alternative hypothesis was that the various phonological processing skills observed in mature language users develop, at least partly, independently. That is, the ability to use phonological structure for linguistic purposes continues to be refined long after a child demonstrates sensitivity to that structure. To meet the goal of the study, serial recall by normal readers of three ages and by poor readers of the youngest age was examined.

Of course, simply exploring whether the use of a phonological coding strategy in working memory is something that develops with age (and so with language experience) could be considered a sufficient goal in itself because only one study has explicitly compared the use of phonological coding strategies in working memory by adults and children. Treiman (1995) tested adults and children of three ages (kindergarten, third grade, and sixth grade) on serial recall of nonsense syllables and observed a developmental trend to greater accuracy in recall. However, the majority of errors observed for all groups were item errors, as would be expected for nonsense items (Treiman & Danis, 1988), and so the question remains open as to whether accuracy in order recall improves with age.

In the current study, we ensured that order recall (rather than item recall) was investigated by using a closed set procedure in which the participants were told the items that would make up the lists before the testing began. The items were consonant–vowel–consonant nonrhyming and rhyming nouns, and the lists were presented auditorily. The participants were asked to arrange pictures depicting the nouns to match the order in which they were heard. Awareness of phonological structure was also evaluated with two separate tasks. The possibility that the use of a phonological coding strategy for storing linguistic items in working memory emerges automatically once an individual has access to phonological structure was examined in two ways. First, group differences in serial recall of rhyming and nonrhyming words were measured for normal and poor readers, who, predictably, demonstrated strong group differences in phonological awareness. A strong group difference should be found in the effect rhyming has on serial recall if the ability to use a phonological coding strategy emerges automatically with phonological awareness. Second, the relation between phonological awareness and serial recall was examined across a range of reading abilities. A strong relation should be found between these two skills if the ability to use a phonological code to store linguistic items in working memory emerges automatically with phonological awareness.

Only one other study has explicitly tried to explore the relation between phonological awareness and serial recall of linguistic materials by children. Mann and Liberman (1984) examined the relation between the ability to count the number of syllables in words and the ability to recall strings of rhyming and nonrhyming words presented auditorily by children in kindergarten and first grade who were classified as "good," "average," or "poor" readers, based on word-attack skills. A Pearson product-moment correlation coefficient calculated between performance on syllable counting and recall for nonrhyming words was

.26. This correlation just reached statistical significance at the .05 level and indicates that phonological awareness (as measured by the syllable counting task) explains about 7% of the variance in recall for nonrhyming words. However, it is the difference in recall of rhyming and nonrhyming words that indexes the extent to which a phonological coding strategy is used for storing items in working memory (i.e., what is commonly termed the "rhyming effect"). Moreover, syllable counting does not evaluate participants' knowledge of syllable-internal structure, and it is specifically this knowledge and its subsequent use for storing items in working memory that would most explain the recall advantage for nonrhyming items commonly observed. In other words, when rhyming effects are observed it presumably means that the phonological code used to store items in working memory is actually a phonemic code. Storing items using a syllabic code should do little to produce a rhyming effect. Even syllables that rhyme are unique because of differences at the first syllable margin. In the end, then, the Mann and Liberman study provides little meaningful insight into the relation between attaining access to phonological structure and the ability to use that structure to store linguistic items in working memory. In the present study, the phonological awareness tasks used with children expressly examined sensitivity to syllable-internal (i.e., phonemic) structure. The assumption was made that it is specifically a phonemic coding strategy that would provide an advantage for the recall of nonrhyming words over rhyming words.

Because the children participating in this study differed in reading ability, care was taken to minimize other task requirements. Several investigators (Bar-Shalom et al., 1993; Smith et al., 1989) have suggested that the abilities of poor readers to hold linguistic information in working memory should improve when processing demands are minimized. Although the specific processing demands these investigators described pertain largely to semantic and pragmatic constraints on sentence comprehension, a similar suggestion could be made for the recall of strings of unrelated words. That is, if processing limitations account for some of the differences previously observed between normal and poor readers in their use of a phonemic coding strategy for serial recall, any procedure that reduces processing demands should reduce the disparity between the performance of normal and poor readers such that the performance of the poor readers improves. Any remaining disparity between the two groups would represent a true difference in the extent to which a phonemic coding strategy is used to store items in working memory. The procedures used in this study were meant to reduce processing demands as well as other task requirements. Pictures were used, and so participants did not need to memorize the items per se. Having participants arrange pictures also meant that this task was more complementary to tasks used in studies of sentence comprehension (i.e., object manipulation and sentence–picture matching) than verbally responding would have been. Not requiring verbal responses was also considered advantageous because it meant that the participants did not need to plan productions. Thus, the possibility was eliminated that the subtle production differences sometimes observed between normal and poor readers (e.g., Brady, Poggie, & Rapala, 1989) would account for some of the differences in order recall between the groups, if found.

EXPERIMENT I: EIGHT-ITEM LISTS

For the first experiment, lists of eight words were used. The decision was made to use lists of this length because many of the early studies on serial recall used lists that were seven to nine items long (e.g., Campbell & Dodd, 1980; Campbell, Dodd, & Brasher, 1983; Darwin & Baddeley, 1974; Spoehr & Corin, 1978). Pilot testing showed that 8-year-olds were able to do this task with lists of this length.

METHOD

Participants

There were four groups of participants: adults, 11-year-olds with normal reading abilities, and two groups of 8-year-old children, defined as either normal or poor readers. All participants passed hearing screenings of the frequencies of 0.5, 1.0, 2.0, 4.0, and 6.0 kHz presented at 25 dB HL. The 17 adults who participated were between the ages of 20 and 40 years, and none had any history of speech or language problems. All the adults passed the reading subtest of the Wide Range Achievement Test–Revised (WRAT-R) (Jastak & Wilkinson, 1984) with a reading ability better than the eleventh grade level.

The 16 11-year-olds who participated had a mean age of 11;6 and were in the middle of sixth grade. None had any history of speech or language problems, and all passed screenings for general nonverbal and language abilities. These children were required to score no more than two standard deviations below the population mean on the block design of the Wechsler Intelligence Scale for Children-III (WISC-III) (Wechsler, 1991) and on the Peabody Picture Vocabulary Test–Revised (PPVT-R) (Dunn & Dunn, 1981). In actuality, means for both tasks were 0.67 standard deviations above the population mean, and the group standard deviations were equal to one population standard deviation. All the children had scores on the reading subtest of the WRAT-R of 95 or better ($M = 114$, $SD = 8$).

A total of 31 children between the ages of 7;9 and 8;11 participated: 20 normal readers and 11 poor readers. The mean age for both groups was 8;3. The 8-year-olds were either close to finishing second grade or had just done so. These children were categorized as normal or poor readers according to their scores on the reading subtest of the WRAT-R. Normal readers were those with standard scores of 95 or better, and poor readers were those with standard scores of 85 or poorer. Specifically, the normal readers had a mean standard score of 105.7 ($SD = 9.4$). Their reading abilities were equivalent to those that are expected near the middle of third grade, putting these children roughly half a year ahead of expectations for their chronological age. The poor readers had a mean standard score of 74.2 ($SD = 11.3$). Their reading abilities were equivalent to those that are expected just past the middle of first grade. Thus, the poor readers' reading abilities were almost two years behind those of the normal readers and roughly a year and a half behind expectations for their chronological age. All the 8-year-olds had normal speech production abilities, as evaluated by the Goldman–Fristoe Test of Articulation–Revised (Goldman & Fristoe, 1986): the normal readers had a mean percentile score of 93 ($SD = 14$), and the poor readers had a mean percentile score of 88 ($SD = 17$). All the 8-year-olds were required to obtain a score better than the 70th standard score (i.e., two standard deviations below the mean) on the PPVT-R. In fact, the normal readers' mean standard score was 99.8 ($SD = 14.6$), and the poor readers' mean standard score was 93.1 ($SD = 11.6$). The Coloured Progressive Matrices (CPM) (Raven, 1975), the children's version of the Raven's Progressive Matrices, was used to screen nonverbal intelligence because it has been used previously in studies with poor readers of this age (e.g., Pratt & Brady, 1988). The normal readers' mean percentile score was 78 ($SD = 24$), and the poor readers' mean percentile score was 70 ($SD = 26$). These differences between the groups (on the PPVT-R and CPM) were not statistically significant.

Equipment

All testing took place in a sound-attenuated booth. Hearing was screened with a Welch Allyn TM262 audiometer/tympanometer with TDH-39 earphones. For the phonological awareness tasks, recorded stimuli were presented with a Nakamichi MR-2 audiocassette player, a Tascam PA-30B amplifier, and a Realistic speaker. For the serial recall task, stimuli were stored on a computer. Both the presentation of stimuli and the recording of responses in this task were controlled by a computer. A Data Translation 2801A digital-to-analog converter, a Frequency Devices 901F analog filter, a Crown D-75 amplifier, and AKG 141 headphones were used to present the stimuli to the participants in this task.

Stimuli

Phonemic awareness. To evaluate phonemic awareness for the 11- and 8-year-olds, two tasks were used, both developed by Pennington and colleagues at the University of Denver (e.g., Pennington et al., 1990): a phoneme deletion task and a pig Latin task. Two phonemic awareness tasks were used because it has been demonstrated that different tasks are needed to distinguish between individuals with good and poor phonemic awareness at different stages of development (Stanovich et al., 1984). The pig Latin task is more difficult than the phoneme deletion task because it requires that a segment not simply be removed, but be moved and recombined with another phonemic unit. Because we were not certain beforehand which task would be most sensitive to the differences in phonemic awareness among children of these ages, we used both tasks. The phoneme deletion task consisted of nonsense syllables that become real words when one segment is removed. There were 6 practice items and 32 test items on this task, and these are listed in Appendix 1. The stimuli for the pig Latin task were real monosyllabic and disyllabic words for which the child was expected to provide the pig Latin form. The one difference between this pig Latin task and the common playground variety is that the children were explicitly trained to move only the first segment of a cluster. For example, the correct pig Latin form of *stick* in this study was *ticksay*. The version of this task used by Pennington and colleagues has 12 practice items and 48 test items (listed in Appendix 2). This version was used with the 11-year-olds. The 8-year-olds were tested on only the first 30 items because pilot work showed that 8-year-olds made numerous errors on this task and could become visibly frustrated. Thus, 30 items seemed to provide the variability needed to capture the differences between normal and poor readers without frustrating any child.

For each task, an audiotape was prepared that provided initial instructions, practice items, and test items. In this way, experimenter variables (such as differences in speaking style or instruction given) that might influence responses were kept to a minimum. For each practice and test item on the phoneme deletion task the following phrase was used: "Say [nonsense syllable]; now say [nonsense syllable] without the [one segment]." For example, "Say *trisk*; now say *trisk* without the /s/. (The sound of the segment is given rather than the letter name.) For the practice and test items on the pig Latin task, individual words were presented in the carrier phrase "The next word is ____."

Serial recall. Two sets of stimuli were prepared to test the serial recall of linguistic materials, one with nonrhyming words and one with rhyming words. All stimuli were consonant–vowel–consonant nouns of which pictures could easily be prepared. The nonrhyming stimuli consisted of the words *teen, ball, coat, pack, dog, ham, rake,* and *seed*. The rhyming stimuli were *mat, bat, gnat, cat, hat, rat, vat*, and *Pat* (Pat was a boy). Three tokens of each word, spoken by a man, were digitized at a 10-kHz sampling rate. Tokens of each word were selected so that the eight words used matched each other closely in terms of fundamental frequency, amplitude, and duration. In addition to these test items, samples of sixteen letters were prepared: eight were nonrhyming (*K, S, R, Q, Y, F, L, H*), and eight were rhyming (*G, D, Z, C, B, T, P, V*). Hand-drawn pictures ($2'' \times 2''$) were prepared to represent each word and each letter.

Procedures

The screening procedures were always administered first. The items for the phonemic awareness tasks were presented via a loudspeaker, calibrated to present the items at a peak intensity of 68 dB SPL measured at the place where the participants sat. The experimenter sat across the table from the participant for the phonemic awareness tasks. Help and/or feedback was provided on practice items only. On both phonemic awareness tasks, testing was stopped after six consecutive errors. The numbers of correct items on each of the two tasks were used in further analysis.

Items for the serial recall task were presented via headphones, calibrated to present items at a peak intensity of 68 dB SPL. The experimenter sat to the side of the participant for the serial recall task. Half the participants in each group heard the nonrhyming lists first, and half heard the rhyming lists first. The computer program that presented stimuli randomized the order of presentation of stimuli separately for each presentation. The experimenter introduced each letter in the letter list to be used for practice (either nonrhyming or rhyming) one at a time and said the letter name as it was introduced. Each letter card was placed on the table facing the experimenter. After all the letter cards were placed on the table, the experimenter demonstrated how to do the task once. That is, the experimenter and participant heard one list of letters, and then the experimenter arranged the letter cards so that they were in the order presented. Several aspects of the procedure were emphasized: the participant's hands had to remain off the table until all list items were heard, and there could be no vocalization until the participant was finished arranging the pictures in the order remembered. The first instruction was given to prevent the children from putting a finger on or near the first item heard, a strategy several children tried to invoke during pilot testing. The second instruction was given to preclude overt rehearsal. It also served to speed up responding, which presumably would have minimized covert rehearsal. The participant was then provided with four practice lists. Correction was provided if the participant had difficulty with the picture arranging procedure itself, but correct ordering of the letters was not required during practice.

Testing occurred next. Before testing with either the rhyming or nonrhyming list, each picture was laid down in front of the participant in turn, and the label was said by the experimenter. When all eight pictures were in front of the participant, the experimenter asked the participant to name each picture. In most cases, the participants could do this easily. However, if a participant misnamed one of the pictures, the experimenter provided the correct name and again asked the participant to name each picture. Testing with the ten lists then took place. The experimenter listened to the presentation of the practice lists and then removed the headphones so that the order of the words presented during testing was not known to her. The experimenter wrote down the order of the pictures after the child arranged them, using the first letter of each word only (to save time). These lists were then compared to the lists of word orders actually presented, which were generated by the program anew for each participant. The number of errors for each list position (out of a total of 10) was computed.

Adults. The screening and serial recall tasks were each presented in the same session. Nothing else was expected of the adults.

11-year-olds. Because correlations were to be computed between the 8-year-olds' scores on the phonemic awareness and serial recall tasks, it was necessary to evaluate the reliability of these tasks. The 11-year-olds provided the data for these measures of reliability. They were selected to provide the data for reliability measures because the tasks were easier for them than for the 8-year-olds and yet they were children. If adults only had provided the data for the reliability measures, the generalizability of those measures to children would be suspect. While the 8-year-olds were able and willing to do both tasks, the tasks were slightly easier for the 11-year-olds. Concern for the use of young children in research suggested that only the data necessary to meet the goals of the study be collected from them. For the phonemic awareness tasks, scores for odd and even items were compared. For the serial recall task, the 11-year-olds were

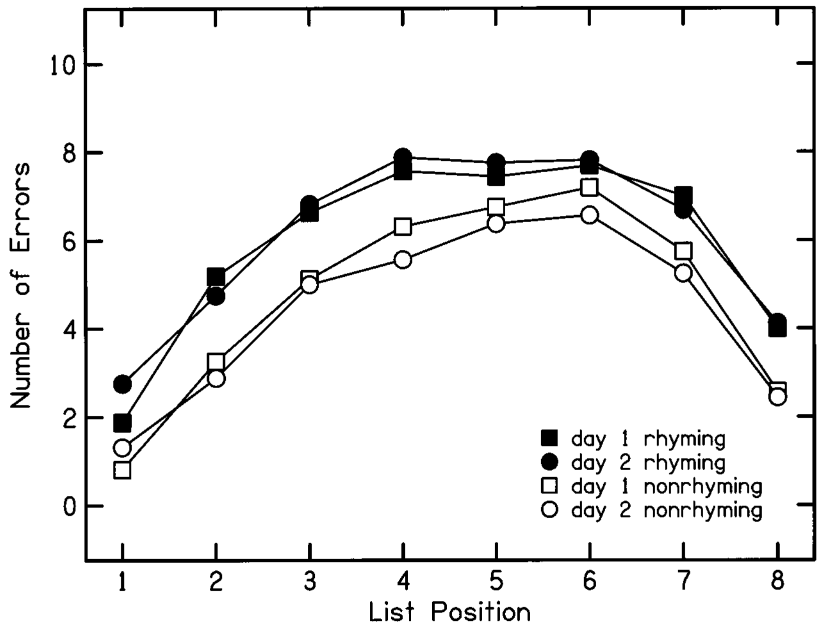

Figure 1. Number of errors made on each day by the 11-year-olds for serial recall of eight-word lists of rhyming and nonrhyming materials.

tested twice on different days, and the scores for these two sessions were compared.

8-year-olds. These children were given the phonemic awareness tasks in the same session as the screening procedures, and the serial recall task was given in a separate session.

RESULTS

Reliability

Reliability coefficients were derived for the phonemic awareness measures and the serial recall measures for the data collected from the 11-year-olds. Mean scores on the phoneme deletion task were 14.6 for the odd items and 14.3 for the even items. Scores on the pig Latin task were 19.3 for the odd items and 18.4 for the even items. The Spearman–Brown formula for estimating reliability from split-halves data was used (e.g., Ferguson, 1981) and yielded reliability coefficients of 0.77 for the phoneme deletion scores and 0.98 for the pig Latin scores. Figure 1 shows mean errors on the serial recall task for each day for rhyming and nonrhyming lists separately. To estimate reliability for the serial recall task, a Pearson product-moment correlation coefficient was computed on

Table 1. *Mean errors (out of 80) across all list positions for the adults, 11-year-olds, and 8-year-olds by rhyme condition and mean difference scores*

	Adults	11-year-olds	8-year-olds: Normal readers	8-year-olds: Poor readers
Rhyming (errors)				
M	43.9	47.4	55.4	59.8
SD	(4.9)	(7.0)	(6.0)	(6.1)
Nonrhyming (errors)				
M	27.8	37.8	47.8	52.4
SD	(6.6)	(8.9)	(10.1)	(6.0)
Difference scores				
M	16.1	9.6	7.6	7.5
SD	(5.8)	(8.7)	(10.3)	(5.5)

the mean number of errors for each day across position and rhyming condition. The resulting reliability coefficient was 0.75. Thus, all of the measures met the criterion of being sufficiently reliable for the early stages of predictive research (Nunnally & Bernstein, 1994).

Developmental trends in serial recall

Data from the first day of testing for the 11-year-olds were used in subsequent analyses. Table 1 displays mean errors across list positions for rhyming and nonrhyming materials separately as well as mean difference scores (i.e., the difference in total errors for rhyming and nonrhyming materials). This last score may be thought of as an operational definition of the rhyming effect. Data for the adults, 11-year-olds, and normal-reading 8-year-olds are shown in columns 1, 2, and 3, respectively, and illustrate developmental trends. Figure 2 shows the mean number of errors at each list position for the three groups of listeners. From Table 1 and Figure 2 it appears that there is a general development improvement in the accuracy of recall, and specifically that the effect of rhyming increases with age. This last developmental trend appears due to older listeners showing a much greater advantage than younger listeners for the nonrhyming materials.

A two-way analysis of variance (ANOVA) was done on the total number of errors across list positions, with age as the between-subjects factor and rhyming as the within-subjects factor. The decision was made to conduct the analysis this way, summing across list positions rather than including list as a factor in the analysis, because it simplified the analysis while preserving the variables of interest. A response was considered wrong if it was given in the incorrect order, and so these summed scores preserved information about the order effect. Preliminary analyses showed that the data were both normally distributed and ho-

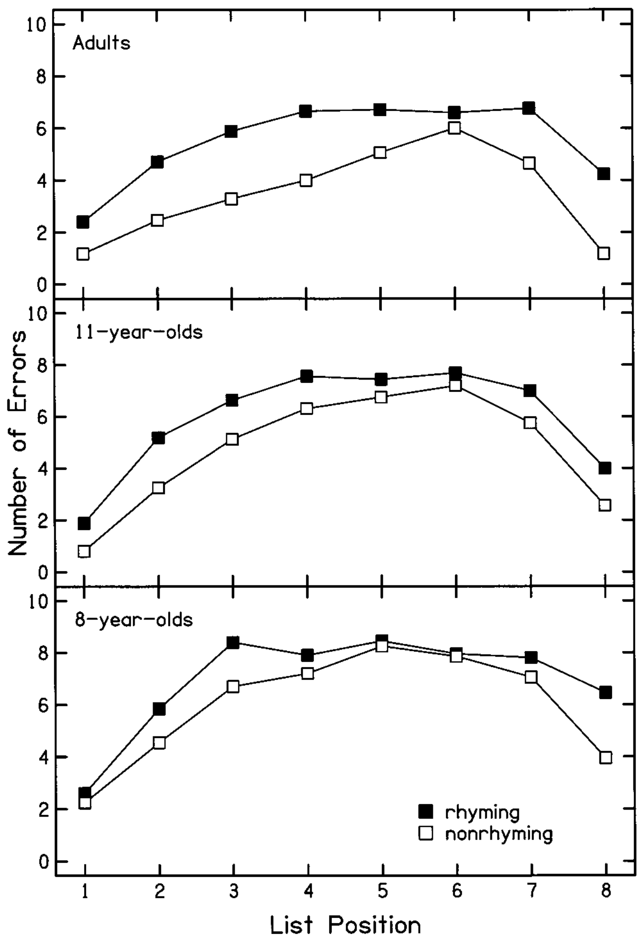

Figure 2. Number of errors made by the adults, 11-year-olds, and normal-reading 8-year-olds for serial recall of eight-word lists of rhyming and nonrhyming materials.

mogeneous with regard to variances across groups. Both main effects were significant: for age, $F(2, 50) = 30.90$, $p < .001$; for the rhyming condition, $F(1, 50) = 87.64$, $p < .001$. Of primary interest, the Age × Rhyme interaction was significant, $F(2, 50) = 4.77$, $p = .013$, indicating that the magnitude of the rhyming effect varied across age.

Next, one-way ANOVAs, with age as the factor, were conducted on scores for the rhyming and nonrhyming lists separately. Pairwise *t* tests with Bonferroni corrections were also done. Results for the rhyming materials showed a significant age effect, $F(2, 50) = 18.09$, $p < .001$, and significant pairwise *t* tests for adults versus 8-year-olds ($p < .001$) and for 11-year-olds versus 8-year-olds ($p < .01$). The *t* test for adults versus 11-year-olds did not reach statistical significance. Results for the nonrhyming materials showed a significant age effect, $F(2, 50) = 24.19$, $p < .001$. All three *t* tests were significant for these materials: adults versus 8-year-olds ($p < .001$), adults versus 11-year-olds ($p < .01$), and 11-year-olds versus 8-year-olds ($p < .01$). Consequently, it seems fair to conclude that there was a general decrease in the number of errors made in serial recall of these materials, both rhyming and nonrhyming, with a stronger effect for the nonrhyming materials. This conclusion is complementary with that from the two-way ANOVA (i.e., the magnitude of the rhyming effect changed with age).

Normal versus poor readers

Phonemic awareness. For the 8-year-olds in the normal-reading group, the mean number of items correct on the phoneme deletion task was 19.6 ($SD = 7.8$). For the children in the poor-reading group, the mean number of items correct was 9.5 ($SD = 4.9$). This group difference was statistically significant, $t(29) = 3.88$, $p < .001$. For the pig Latin task, the mean number of items correct for the normal-reading group was 15.5 ($SD = 11.3$). The mean number of items correct for the poor-reading group was 1.5 ($SD = 3.9$). This group difference was statistically significant, $t(29) = 3.97$, $p < .001$.

Serial recall. The last two columns of Table 1 provide mean error scores for the normal-reading and poor-reading 8-year-olds. Figure 3 shows the mean number of errors across participants in each reading group for each list position for nonrhyming and rhyming materials. A two-way ANOVA was performed on the summed error scores across list positions, with reading ability as the between-subjects factor and rhyming condition as the within-subjects factor. Only the main effect of rhyming condition was statistically significant, $F(1, 29) = 20.02$, $p < .001$. The effect of reading ability did not quite reach statistical significance, $F(1, 29) = 3.84$, $p = .059$. Of most interest, the Reading Ability $\times$ Rhyme interaction was not significant.

Correlations. Several Pearson product-moment correlation coefficients (*r*) were computed to see if either of the phonemic awareness measures or the serial recall difference score were significantly related to reading ability, as measured by the WRAT-R, and to see if the serial recall difference score and the phonemic awareness measures were related to each other. Scores for both phonemic awareness tasks were correlated with scores on the reading subtest of the WRAT-R: for phoneme deletion ($r = .71$, $p < .001$), for pig Latin ($r = .54$, $p = .002$). This latter correlation may have been weaker than that for the phoneme deletion task because so many of the poor readers were simply unable to do the pig Latin task. The correlation coefficient for the serial recall difference scores and the scores on the reading subtest of the WRAT-R was not significant. Although there was scant evidence from the group data that the normal readers were using a phonemic coding strategy to a greater extent than the poor readers

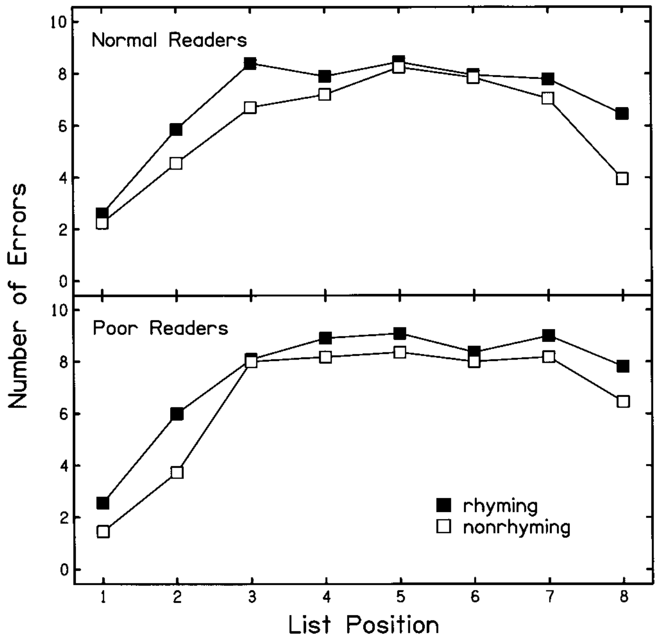

Figure 3. Number of errors made by the good and poor readers for serial recall of eight-word lists of rhyming and nonrhyming materials.

for storing items in working memory, it still seemed worthwhile to investigate whether phonemic awareness accounted for any of the variance in the serial recall difference scores. To this end, a Pearson product-moment correlation coefficient was computed between scores on each of the phonemic awareness tasks and the serial recall difference scores. If phonemic awareness accounted for any portion of the variance in serial recall for linguistic materials, these correlations should be significant. Neither one was.

DISCUSSION

This study showed a clear developmental trend in the use of phonemic codes for storing linguistic items in working memory. Children as old as 11 years made significantly more errors on the nonrhyming materials than the adults made, indicating that they were not coding items in memory with a phonemic code as strongly as the adults were. Evidence of phonemic coding strategies was even weaker for the 8-year-olds.

Another finding of this study was an observed dissociation between phonemic awareness and phonemic coding of linguistic materials in working memory. The normal readers in this study demonstrated significantly greater skill at manipulating the phonemic structure of syllables than the poor readers, indicating that they had access to that structure. Nonetheless, the results for the serial recall task for the normal readers showed no evidence that a phonemic coding strategy had been used to a greater extent by them than by the poor readers to store words in working memory. Several results support this conclusion. The normal readers made somewhat fewer errors overall than the poor readers (although the effect did not reach statistical significance), but they showed no more of a rhyming effect than the poor readers showed: the differences in the number of errors made between the normal and poor readers were roughly equal for both rhyming and nonrhyming lists. If observed differences between the reading groups were attributable to differences in the extent to which a phonemic code was used to store words in working memory, the largest group difference would have been observed for nonrhyming materials, as was the case for the adults and the 11-year-olds: mean number of errors were similar for rhyming materials for the adults and the 11-year-olds but significantly different for nonrhyming materials. Also, the relation between phonemic awareness (as measured by both the phoneme deletion and pig Latin tasks) and the serial recall difference score was not statistically significant. Thus, no group difference was observed in the extent to which phonological coding strategies were used, nor was there a relation between phonemic awareness and the use of phonological coding strategies across the range of reading abilities. It seems that the use of a phonemic coding strategy for storing items in working memory does not follow automatically as a result of developing access to the syllable-internal, phonemic structure of language. Instead, it seems that learning to store items in working memory using a phonemic code takes place over time and developmentally trails the ability to retrieve phonemic information from the linguistic signal.

Care was taken in this study to minimize task requirements. Nonetheless, these results showed little disparity between the performance of the normal and poor readers on the recall task. Therefore, the conclusion could be reached that the poor readers' serial recall improved in this study, relative to that of earlier studies showing differences in serial recall between normal and poor readers, such that they performed similarly to the normal readers. While minimizing task requirements may have accounted for the decrease in disparity between the normal- and poor-reading 8-year-olds to some extent, neither group performed as well as the 11-year-olds or adults. Thus, it seems fair to suggest that some additional skill must be needed besides being able to access phonemic structure in order to make use of that structure in working memory.

At the same time, the serial recall task used in this experiment may have been too difficult generally for 8-year-olds, thus degrading the performance of the normal-reading 8-year-olds. It was not uncommon for 8-year-olds, even those with normal reading abilities, to make 10 errors on some items in the intermediate list positions in the serial recall task, as suggested by the high mean error rates for these positions seen in Figures 1 and 2. In addition, some of the 8-year-olds were unable to obtain any correct answers on the pig Latin task. Therefore, our ability to detect significant group differences and significant correlations may have been constrained. Experiment 2 was designed to see if the results of this first experiment would be replicated when these potential problems were corrected.

EXPERIMENT II: SIX-ITEM LISTS

A second experiment was conducted as a check on the possibility that the failure to find the anticipated effects in Experiment 1 for normal and poor readers was due to what is traditionally termed ceiling and floor effects. Specifically, a significant Reading Ability × Rhyme interaction was expected for the serial recall task but was not found. In addition, significant correlations were expected both between the reading scores and the serial recall difference scores and between the phonemic awareness scores and the serial recall difference scores. None of those expected correlations was observed. In this second experiment, several procedural changes were made. First, the serial recall task was conducted with six-item lists in hopes that any ceiling effects for the numbers of errors would be avoided. Also, all 48 items were used on the pig Latin task in hopes that floor effects for the numbers of items correct on that task would be avoided. This second change would not be expected to have much effect on its own for the 8-year-olds who simply could not do the pig Latin task: those children did not get further than the first six items. Nonetheless, it seemed a worthwhile attempt to spread out scores on the pig Latin task. Finally, the age range of children included in the comparison of good and poor readers was increased to include 8-, 9-, and 10-year-olds. In addition to improving the chances that most of the children would not score near the floor on the pig Latin task, this change meant that the number of children participating would be increased, thus improving the possibility of finding significant group differences and correlations, if they actually exist in the general population.

METHOD

Participants

Children between 8 and 10 years of age were enlisted for this second experiment. Only one change was made to the criteria for participation from the first experiment. The block design of the WISC-III was used to screen the children for nonverbal abilities instead of the CPM. This change was made because the mean scores for the 8-year-olds in the first experiment were higher than would be expected for a randomly selected group of children, if the test norms were appropriate for these samples of children. A total of 73 children met the criteria for participation. Of these, 57 children fit the description of normal readers (standard scores for the reading subtest of the WRAT-R of 95 or better) and 16 children fit the description of poor readers (standard scores of 85 or poorer) used in the first experiment. Mean age of the participants in each group was 9;3. Mean standard scores on the reading subtest of the WRAT-R were 108 for the normal readers ($SD = 7$) and 76 for the poor readers ($SD = 9$). As in the first experiment, these scores meant that the normal readers were reading roughly half a year above expectations for their chronological age, and the poor readers were reading roughly a year and a half behind expectations for their chronological age. Unlike the first experiment, though, slight differences were found between the normal- and poor-reading groups on the criterion measures of general and language abilities. On the block design of the WISC-III, the mean score for normal readers was .33 standard deviations above the mean, whereas the mean score for poor readers was .33 standard deviations below the mean. Within-group standard deviations were the same as in the general population for both groups. The between-group difference was statistically significant, $t(71) = 2.26$, $p = .03$. On the PPVT-R, the mean standard score for the normal readers was 106 ($SD = 13$), and the mean score for the poor readers was 93 ($SD = 11$). This difference was also statistically significant, $t(71) = 3.65$, $p < .001$. These group differences were not considered problematic in this experiment, largely because they were actually small in magnitude: group means for both groups were very close to the population means. Furthermore, such differences could only increase the probability of finding group differences on the recall task, and the prediction in this experiment was that no such differences would be found. Specifically, the effects of interest were the Reading Ability × Rhyme interaction, the correlation between serial recall difference scores and reading scores, and the correlations between each of the phonemic awareness measures and serial recall difference scores. Failure to find these effects statistically significant, even though slight differences in general and language abilities exist, would only provide particularly strong support for the contention that the effects do not exist in the general population.

Stimuli and procedures

With two exceptions, the stimuli and procedures were the same as in Experiment 1. First, all 48 items were used in the pig Latin task. Second, the lists of words for the serial recall task consisted of six items instead of eight. For the nonrhyming lists, the words *teen* and *seed* were excluded; for the rhyming lists, the words *Pat* and *vat* were excluded. For the training lists, the letters *K* and *L* were excluded from the nonrhyming lists, and *G* and *B* were excluded from the rhyming lists.

RESULTS

Phonemic awareness

For the normal-reading group, the mean number of items correct on the phoneme deletion task was 23.9 ($SD = 6.3$). For children in the poor-reading group, the mean number of items correct was 13.4 ($SD = 7.2$). This group difference was statistically significant, $t(71) = 5.71$, $p < .001$. For the pig Latin task, the mean number of items correct for the normal-reading group was 29.3 ($SD = 14.1$). The mean number of items correct for the poor-reading group was 9.3 ($SD = 12.3$). This group difference was statistically significant, $t(71) = 5.16$, $p < .001$.

Serial recall

Table 2 shows mean error scores for the normal and poor readers as well as the mean difference scores. As in Experiment 1, these scores are summed across list positions but provided separately for the rhyming and nonrhyming lists.

Table 2. *Mean errors (out of 60) across all list positions for the normal and poor readers by rhyme condition and mean difference scores*

	Normal readers	Poor readers
Rhyming (errors)		
M	30.3	34.8
SD	(7.8)	(5.0)
Nonrhyming (errors)		
M	24.2	29.4
SD	(8.8)	(6.6)
Difference scores		
M	6.1	5.3
SD	(7.4)	(6.0)

Figure 4 shows the mean number of errors for each list position. A two-way ANOVA was performed on the summed error scores across list positions, with reading ability as the between-subjects factor and rhyming condition as the within-subjects factor. As in Experiment 1, the main effect of rhyming condition was statistically significant, $F(1, 71) = 32.53$, $p < .001$. This time, the effect of reading ability was clearly significant, $F(1, 71) = 5.99$, $p = .017$. Again, however, the Reading Ability × Rhyme interaction was not significant.

Correlations

The same Pearson product-moment correlation coefficients were computed on these data as on those of Experiment 1, with the same results. The correlations between scores on each phonemic awareness task with scores on the reading subtest of the WRAT-R were statistically significant: for phoneme deletion ($r = .70$, $p < .001$), for pig Latin ($r = .57$, $p < .001$). These correlations are strikingly similar to those computed for the data in Experiment 1. Thus, even with more items on the pig Latin task and a wider range of participant ages, the correlation between this phonemic awareness task and reading ability was not as great as between phoneme deletion and reading ability. As in Experiment 1, the correlation between the serial recall difference scores and the scores on the reading subtest of the WRAT-R was not significant. Finally, the correlations computed between scores on each phonemic awareness measure and the serial recall difference scores were not significant.

DISCUSSION

The purpose of the second experiment was to check the findings from Experiment 1 to ensure that observed effects were not attributable to procedural limitations. In spite of the changes in procedures, identical trends were observed in the second experiment: the poor readers made somewhat more errors on serial

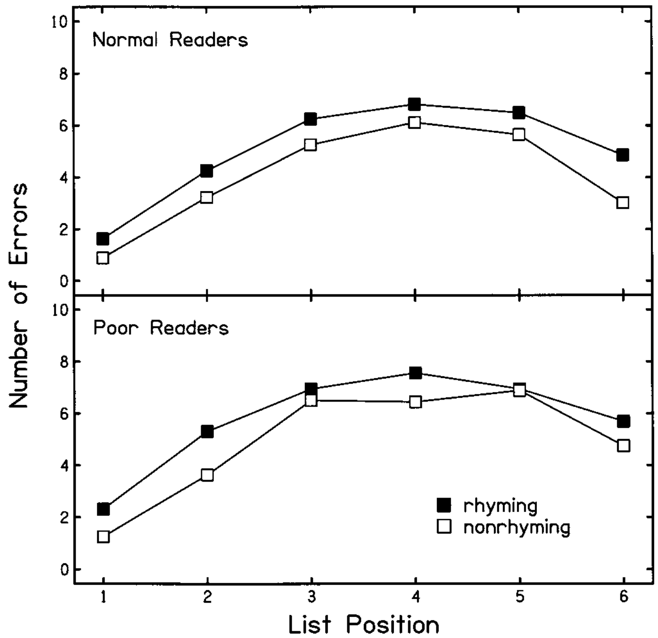

Figure 4. Number of errors made by the good and poor readers for serial recall of six-word lists of rhyming and nonrhyming materials.

recall than the good readers, but the differences between the groups were similar for the rhyming and nonrhyming materials. Consequently, no evidence was found that good readers use a phonemic code to a greater extent than poor readers. In addition, the serial recall difference score was found to correlate neither with reading ability nor with phonemic awareness.

GENERAL DISCUSSION

A central question addressed by this study was whether children who are capable of accessing phonemic structure in the speech signal necessarily use that structure for storing items in working memory to a greater extent than children who have difficulty accessing phonemic structure. In other words, does the use of phonemic coding strategies in working memory emerge automatically with the ability to access that phonemic structure? Based on the results of two experiments, the answer to this question is apparently "no." Young normal readers showed no evidence of using phonemic coding strategies for storing items in working memory to a greater extent than young poor readers. First, no differences were found between the reading groups in the magnitude of the rhyming effect. Second, no relation was found between the children's reading abilities and the extent to which they used a phonemic code for storing items in working memory.

A second question addressed by this study was whether there is a developmental change in the extent to which phonemic coding strategies are used for storing items in working memory. As obvious as this question seems, there has been little direct investigation of it. Based on the results of the first experiment, the answer to this question is a clear "yes." Interestingly, children as old as 11 years who were able to access syllable-internal phonemic structure did not use that structure to store words in working memory to the same extent that the adults did. Instead, a developmental enhancement in the use of phonemic coding strategies for storing words in working memory was observed. The ability to use syllable-internal phonemic structure in further linguistic processing apparently continues to emerge long after a child has acquired the ability to access that structure.

Of course, the question now arises as to why, if they are capable of accessing phonemic structure, do children with normal reading abilities fail to make use of phonemic coding strategies for storing words in working memory? Two different explanations can be imagined. First, placing linguistic materials into working memory with a phonemic code may be a processing step separate from the extraction of phonemic structure in words. That is, the normal readers may have been recovering phonemic structure from the linguistic signals presented in these experiments but may have been unable to use that information for placing words into working memory. A second possible reason that the normal-reading children may have failed to use phonemic codes for storing words in working memory is that in this situation they may have, in fact, lacked access to phonemic structure. In their study, Spring and Perry (1983) examined digit naming speed and serial recall by adequate and poor readers. A significant correlation was found between digit naming speed and serial recall of nonrhyming words.[2] Thus, processing speed is apparently a factor in the extraction of phonemic structure and/or phonemic coding for serial recall. The normal readers in the present study may simply have been unable to extract phonemic structure from the stimuli quickly enough to take full advantage of a phonemic coding strategy. This possibility was not eliminated by the phonemic awareness tasks used here because our participants were not required to respond quickly to those tasks. Consequently, the children had the time they needed to access and then manipulate the phonemic structure of the words. Future experiments will need to resolve the question of why children with good awareness of phonemic structure fail to use that structure to store linguistic materials in working memory.

Of course, the 8-year-olds classified as normal readers did make fewer errors overall than the 8-year-olds classified as poor readers. Thus we may ask how the good readers were storing words in working memory if not with a phonemic code. A clue may come from the studies of Montgomery (1995b) and Mann and Liberman (1984). Neither the nonsense repetition task of Montgomery nor the syllable counting task of Mann and Liberman necessarily required access specifically to *phonemic* structure. Nonetheless, the children with normal language and normal reading abilities performed better on these tasks than the children with language and/or reading difficulties. Moreover, nonsense repetition correlated with sentence comprehension in the Montgomery study and syllable counting correlated with serial recall of nonrhyming words in the Mann and Liberman study. Thus, the children in those studies who had better access to phonological structure were better able to store long word strings in working memory. It may be that, for those children, words were stored in working memory with a phonological code but not explicitly with a phonemic code.

The results of this study provide some information for theories of syntactic processing. Although these results did not support the hypothesis that children with better access to phonemic structure used that structure to store linguistic items in working memory, those children were nonetheless able to retain words in working memory better than children with poor access to phonemic structure. Therefore, these results provide some support for the notion that differences in storage capacity for linguistic materials account for differences between good and poor readers on tasks of syntactic comprehension (rather than differences in syntactic abilities per se). The unanswered question is, how do children with normal reading abilities store items in working memory if not with an explicitly phonemic code? This question is particularly intriguing, given the evidence that adults do, in fact, use an explicitly phonemic code to store linguistic items in working memory.

In conclusion, this study has demonstrated that the use of a phonemic coding strategy to store linguistic materials in working memory is a skill that emerges separately and developmentally later than the ability to access phonemic structure. The contradictory results emerging from earlier studies of serial recall by children varying in reading ability may reflect slight differences in participant groups used in those studies. Many of the studies reviewed in this article included children roughly 7 to 9 years of age but differed in the range of reading abilities included. Some studies compared recall abilities of children classified only as normal or poor readers, while other studies included children classified as superior readers. Children considered superior readers at these ages may be developmentally advanced enough that they have acquired the ability to use phonemic coding strategies to store linguistic materials in working memory. At any rate, the way in which young children who are developing language normally store words in working memory remains an unresolved question. However, it is a question deserving of more study.

APPENDIX 1
Items from the phoneme deletion task

Practice	
pin(t)	p(r)ot
(t)ink	no(s)te
bar(p)	s(k)elf

Test	
1. (b)ice	17. s(t)ip
2. toe(b)	18. fli(m)p
3. (p)ate	19. c(l)art
4. as(p)	20. (b)rock
5. (b)arch	21. cream(p)
6. tea(p)	22. hi(f)t
7. (k)elm	23. dril(k)
8. blue(t)	24. mee(s)t
9. jar(l)	25. (s)want
10. s(k)ad	26. p(l)ost
11. hil(p)	27. her(m)
12. c(r)oal	28. (f)rip
13. (g)lamp	29. tri(s)ck
14. ma(k)t	30. star(p)
15. s(p)alt	31. fla(k)t
16. (p)ran	32. (s)part

Note: The segment to be deleted is in parentheses. The correct response is apparent by removing the segment to be deleted.

APPENDIX 2
Items from the pig Latin task

Practice			
go	(ogay)	stick	(ticksay)
pat	(atpay)	drip	(ripday)
happy	(appyhay)	strap	(trapsay)
candy	(andycay)	scram	(cramsay)
thick	(ickthay)	snapshot	(napshotsay)
where	(rewhay)	shop	(opshay)
Test			
1. day	(ayday)	25. dragon	(ragonday)
2. box	(oxbay)	26. sprint	(printsay)
3. lady	(adylay)	27. screamer	(creamersay)
4. funny	(unnyfay)	28. game	(amegay)
5. chatter	(atterchay)	29. rabbit	(abbitray)
6. strike	(trikesay)	30. dresser	(resserday)
7. strangle	(tranglesay)	31. mitten	(ittenmay)
8. gray	(raygay)	32. splitting	(plittingsay)
9. third	(irdthay)	33. man	(anmay)
10. happen	(appenhay)	34. choppy	(oppychay)
11. screw	(crewsay)	35. braver	(raverbay)
12. flatter	(latterfay)	36. what	(atwhay)
13. shelter	(eltershay)	37. wind	(indway)
14. steak	(teaksay)	38. fault	(aultfay)
15. shone	(oneshay)	39. green	(reengay)
16. shudder	(uddershay)	40. chicken	(ickenchay)
17. blow	(lowbay)	41. splatter	(plattersay)
18. shiny	(inyshay)	42. thirst	(irstthay)
19. that	(atthay)	43. scratch	(cratchsay)
20. shelf	(elfshay)	44. stronger	(trongersay)
21. strict	(trictsay)	45. blanket	(lanketbay)
22. brief	(riefbay)	46. straw	(trawsay)
23. closet	(losetcay)	47. weather	(eatherway)
24. blend	(lendbay)	48. strainer	(trainersay)

Note: The correct response is given in parentheses.

ACKNOWLEDGMENTS

This work was supported by research grants R01 DC00633 and P60 DC00982 from the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health. The comments on early drafts of this manuscript by Carol A. Fowler, Donna L. Neff, Bruce F. Pennington, and Michael Studdert-Kennedy are greatly appreciated.

NOTES

1. In many studies of the effects of phonological confusibility on serial recall, results are described as demonstrating a penalty imposed on recall by rhyming materials. Here this description is modified to suggest that an advantage is obtained for non-

rhyming materials. Because good and poor readers generally show similar recall accuracy for rhyming materials whereas good readers show greater accuracy than poor readers for nonrhyming materials, we suggest that a benefit is accrued over normal memory limitations when a perceiver can take advantage of phonological factors in the stimuli. Thus, describing the effect as an advantage for nonrhyming materials rather than as a penalty for rhyming materials seems to denote more accurately the direction of the effect.

2. Of course, the Spring and Perry (1983) study is subject to the same constraints as that of Mann and Liberman (1984): serial recall for nonrhyming words alone does not provide an index of the extent to which a phonemic code was used to store those words in working memory; it is the difference in serial recall for rhyming and nonrhyming words. Nonetheless, this result is relevant here because it does suggest that processing speed may be a factor in these tasks.

REFERENCES

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362–365.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 575–589.

Bar-Shalom, E. G., Crain, S., & Shankweiler, D. P. (1993). A comparison of comprehension and production abilities of good and poor readers. *Applied Psycholinguistics, 14,* 197–227.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35,* 345–367.

Brady, S., Poggie, E., & Rapala, M. M. (1989). Speech repetition abilities in children who differ in reading skill. *Language and Speech, 32,* 109–122.

Byrne, B. (1981). Deficient syntactic control in poor readers: Is a weak phonetic memory code responsible? *Applied Psycholinguistics, 2,* 201–212.

Campbell, R., & Dodd, B. (1980). Hearing by eye. *Quarterly Journal of Experimental Psychology, 32,* 85–99.

Campbell, R., Dodd, B., & Brasher, J. (1983). The sources of visual recency: Movement and language in serial recall. *Quarterly Journal of Experimental Psychology, 40,* 571–587.

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

Crain, S. (1989). Why poor readers misunderstand spoken sentences. In D. Shankweiler & I. Y. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 133–165). Ann Arbor: University of Michigan Press.

Crain, S., & Shankweiler, D. (1991). Modularity and learning to read. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception* (pp. 375–392). Hillsdale, NJ: Erlbaum.

Crain, S., Shankweiler, D., Macaruso, P., & Bar-Shalom, E. (1990). Working memory and sentence comprehension: Investigations of children with reading disorder. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 477–508). Cambridge: Cambridge University Press.

Darwin, C. J., & Baddeley, A. D. (1974). Acoustic memory and perception of speech. *Cognitive Psychology, 6,* 41–60.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test–Revised.* Circle Pines, MN: American Guidance Service.

Ferguson, G. A. (1981). *Statistical analysis in psychology and education.* New York: McGraw-Hill.

Goldman, R., & Fristoe, M. (1986). *Goldman–Fristoe Test of Articulation.* Circle Pines, MN: American Guidance Service.

Hall, J. W., Wilson, K. P., Humphreys, M. S., Tinzmann, M. B., & Bowyer, P. M. (1983). Phonemic-similarity effects in good vs. poor readers. *Memory and Cognition, 11,* 520–527.

Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test–Revised.* Wilmington, DE: Jastak Associates.

Katz, R. B., Shankweiler, D. P., & Liberman, I. Y. (1981). Memory for item order and phonetic recoding in the beginning reader. *Journal of Experimental Child Psychology, 32,* 474–484.
Mann, V. A., Cowin, E., & Schoenheimer, J. (1989). Phonological processing, language comprehension, and reading ability. *Journal of Learning Disabilities, 22,* 76–89.
Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities, 17,* 592–599.
Montgomery, J. W. (1995a). Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics, 16,* 355–378.
(1995b). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research, 38,* 187–199.
Nunnally, J. M., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
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.
Pratt, A. C., & Brady, S. (1988). Relation of phonological awareness to reading disability in children and adults. *Journal of Educational Psychology, 80,* 319–323.
Raven, J. C. (1975). *Coloured Progressive Matrices.* London: H. K. Lewis.
Salame, P., & Baddeley, A. (1986). Phonological factors in STM: Similarity and the unattended speech effect. *Bulletin of the Psychonomic Society, 24,* 263–265.
Shankweiler, D. P., Liberman, I. Y., Mark, L. S., Fowler, C. A., & Fischer, F. W. (1979). The speech code and learning to read. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 531–545.
Smith, S. T., Macaruso, P., Shankweiler, D., & Crain, S. (1989). Syntactic comprehension in young poor readers. *Applied Psycholinguistics, 10,* 429–454.
Smith, S. T., Mann, V. A., & Shankweiler, D. (1986). Spoken language comprehension by good and poor readers: A study with the Token Test. *Cortex, 22,* 627–632.
Spoehr, K. T., & Corin, W. J. (1978). The stimulus suffix effect as a memory coding phenomenon. *Memory and Cognition, 6,* 583–589.
Spring, C., & Perry, L. (1983). Naming speed and serial recall in poor and adequate readers. *Contemporary Educational Psychology, 8,* 141–145.
Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38,* 175–190.
Stein, C. L., Cairns, H. S., & Zurif, E. B. (1984). Sentence comprehension limitations related to syntactic deficits in reading-disabled children. *Applied Psycholinguistics, 5,* 305–322.
Treiman, R. (1995). Errors in short-term memory for speech: A developmental study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1197–1206.
Treiman, R., & Danis, C. (1988). Short-term memory errors for spoken syllables are affected by the linguistic structure of the syllables. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 145–152.
Vogel, S. A. (1975). *Syntactic abilities in normal and dyslexic children.* Baltimore, MD: University Park Press.
Wagner, R. K. (1986). Phonological processing abilities and reading: Implications for disabled readers. *Journal of Learning Disabilities, 19,* 623–630.
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.
Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third Edition.* San Antonio, TX: Harcourt Brace Jovanovich.