CHILDREN'S MEMORY FOR SENTENCES AND WORD STRINGS IN RELATION TO READING ABILITY*

Virginia A. Mann, Isabelle Y. Liberman+ and Donald Shankweiler+

Abstract. This study explores earlier indications that ability to make effective use of speech coding in working memory is correlated with success in learning to read. A previous study of recall of letter strings in good and poor beginning readers (Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977) revealed that the performance of good readers was more severely penalized than that of poor readers when the letter names rhymed. In order to determine whether the differences in susceptibility to phonetic interference extend to materials that more closely resemble actual text, we designed an experiment to test recall of phonetically controlled sentences and word strings. As in the case of letter recall, we found an interaction between reading ability and the effect of phonetic confusability: Though good readers made fewer errors than poor readers when sentences or word strings contained no rhyming words, they made as many errors as poor readers when many rhyming words were present. In contrast to the effectiveness of manipulations of phonetic content, systematic manipulations of meaningfulness and variations in syntactic structure did not differentially affect the two reading groups. We conclude that the inferior performance of poor readers in recall of phonetically nonconfusable sentences, word strings, and letter strings reflects failure to make full use of phonetic coding in working memory.

INTRODUCTION

Much evidence suggests that adult subjects employ a phonetic representation during comprehension of both spoken and written material (see, for

* A revised version of this manuscript will appear in Memory & Cognition.
+ Also at University of Connecticut, Storrs.

Acknowledgment: This work was partially supported by a grant from the National Institute of Child Health and Human Development (Grant HD-01994) and by a NICHD Postdoctoral Fellowship (Fellowship HD05677) to Virginia Mann. Thanks are due to Robert Katz who also aided in testing reading ability and IQ. Bruno Repp and Leonard Katz provided helpful assistance with the statistical treatment of the data. We also wish to express our appreciation for the generous cooperation of the staff of the Parker Memorial School in Tolland, Connecticut. Special thanks are due to the principal, Robert Lincoln, the three second grade teachers, Judy Lombard, Dolores Iadoresta and Jean Medick, and the school secretary, Florence Peacock. We are also grateful to the children who participated and to the parents, without whose permission we could not have proceeded.

[HASKINS LABORATORIES: Status Report on Speech Research SR-59/60 (1979)]
example, Baddeley, 1978; Kleiman, 1975; Levy, 1977; Liberman, Mattingly, & Turvey, 1972; Tzeng, Hung, & Wang, 1977). In several studies of beginning readers, we and other investigators (Byrne & Shea, 1979; Liberman, Shankweiler, Liberman, Fowler, & Fischer, 1977; Mark, Shankweiler, Liberman, & Fowler, 1977) have found new support for the involvement of phonetic representation in the reading process. The ability to make effective use of phonetic representation appears to be correlated with success at learning to read.

The possibility of an association between children's reading ability and their use of phonetic representation was first explored by Liberman, Shankweiler and their colleagues (Liberman et al., 1977) who assessed the role of phonetic representation in memory for letter strings. Using a modification of Conrad's (1964) procedure, they asked good and poor readers in the second grade to recall a string of consonants in which the letter names either rhymed or did not. In both the rhyming and nonrhyming conditions, good readers recalled more items than poor readers. However, the good readers, like Conrad's adult subjects, were greatly penalized by rhyme, whereas the poor readers performed at about the same level on both rhyming and nonrhyming strings. A subsequent experiment (Shankweiler & Liberman, 1976) showed that the same pattern of recall performance occurred whether the items were presented by ear or by eye. The interaction of reading ability and the effect of phonetic confusability has also been demonstrated in the case of recognition memory for isolated words, where good readers show evidence of greater reliance on the use of phonetic representation as a means of remembering words presented in either written (Mark et al., 1977) or spoken (Byrne & Shea, 1979) form. From all of these findings it would seem that underlying the defective performance of poor readers is a problem that extends beyond the act of recoding from print to speech, involving a more general deficit in the use of phonetic coding in working memory.

Consonant strings and isolated words, however, are far removed from actual text. It remains to be determined whether good and poor readers' recall of more natural linguistic stimuli will be affected by the same variables that affect the recall of letters and words. Accordingly, in this investigation we have extended our study of the effect of phonetic confusability to the more ecologically valid situation of recall of sentences.

Previous findings in the research literature lead us to expect that poor readers' recall of both sentences (Mattis, French, & Rapin, 1975; Perfetti & Goldman, 1976; Pike, Note 2; Weinstein & Rabinovitch, 1971; Wiig & Roach, 1975) and word strings (Bauer, 1977; Katz & Deutsch, 1964) will be inferior to that of good readers. We might suppose from the results obtained in the case of letter strings (Liberman et al., 1977), that the introduction of phonetically confusable words into the sentence or word string to be recalled would differentially affect children who differ in reading ability. We therefore assessed children's ability to recall sentences which vary not only along the traditional dimensions of syntax and meaning (as in Miller & Isard, 1964), but also in the presence of phonetically confusable words. Our materials included seven different syntactic constructions, each of which is presented in four versions: a meaningful version in which none of the words rhyme, a meaningful version in which the majority of words rhyme, a meaningless version in which the words do not rhyme, and a meaningless version in which most words again
rhyme. The recall of word strings is examined in an analogous fashion, with items containing five words selected from the meaningless versions of the test sentences. In half of these the words do not rhyme; in half, they do rhyme.

METHOD

Subjects

The subjects were second grade children from a public school in suburban Connecticut. An initial subject pool of 15 good and 15 poor readers was obtained by means of teacher recommendations and scores on the word recognition subtest of the Comprehensive Test of Basic Skills (1974), which had been administered at the end of the first grade. The reading ability of subjects selected in this way was assessed by administration of the Word Attack and Word Identification subtests of the Woodcock Reading Mastery Tests (Woodcock, 1973). The mean sum of raw scores on these subtests was 54.2 for good readers, as compared to 133.9 for poor readers, t(28)=18.19, p<.001. There was no overlap between scores of the two groups. The subjects had IQ scores ranging between 90 and 135 on the Slosson Intelligence Test (Slosson, 1963). The mean IQ score for good readers (114.7) was marginally superior to that of poor readers (107.6), t(28)=1.6, p<.06. The two groups were not significantly different in mean age: 96.3 months for the good readers, 97.1 for the poor readers. All children had been screened by the school system and found to be free from speech or hearing disorders.

Materials

Sentences. Items for the sentence repetition task were permutations of seven 13-word sentences of English. These seven base sentences were chosen to represent a variety of English constructions with complexity varied along a number of syntactic dimensions. The adoption of 13 words as the sentence length was motivated by a desire to prevent good readers from achieving ceiling performance, since ceiling performance confounds interpretation of many previous studies of the sentence recall of good and poor readers. Before designing the sentence repetition materials, we conducted a pilot study of the effect of sentence length on the sentence recall of eight average readers in the second grade classrooms from which the subjects were drawn. Results indicated that average readers begin to make errors as the length of a meaningful, phonetically nonconfusable sentence approaches 13 words.

Each sentence was presented in four versions, which were constructed by substitutions among content words with position and choice of function words held constant. Thus syntactic structure was the same across the four versions of each base, while manipulations of content words permitted orthogonal variation of sentence meaning and phonetic confusability. Versions were either meaningful, phonetically nonconfusable; meaningful, phonetically confusable; meaningless, phonetically nonconfusable; or meaningless, phonetically confusable. A representative example of a base sentence and its four versions is given in Table 1.
Table 1

BASE SENTENCE:
/NOUN/’s /ADJ/ /NOUN/ /VERB/ (past tense) at the /NOUN/ that /VERB/ (past tense) on the /ADJ/ /NOUN/.

Versions:
Meaningful, phonetically nonconfusable:

Peg’s brown dog bit at the bone that fell on the clear floor.

Meaningful, phonetically confusable:

Pat’s bad cat bat at the rat that sat on the flat mat.

Meaningless, phonetically nonconfusable:

Bob’s fried cap laughed at the chair that stood on the smart glass.

Meaningless, phonetically confusable:

Kay’s gray hay stayed at the clay that lay on the gay day.

All versions of each base sentence were matched with respect to word frequency (Thorndike & Lorge, 1944) and the number of syllables contained in each word. The meaningless versions differed from meaningful ones with respect to whether choice of nouns, verbs and adjectives adhered to semantic restrictions. Meaningful versions were created in accordance with these restrictions; meaningless versions were created by violating them. The phonetically nonconfusable and phonetically confusable versions differed with respect to the presence of rhyming items. Phonetically nonconfusable versions contained no rhyming words, phonetically confusable versions contained from seven to nine rhyming words. The number of rhyming words and their position were held constant across the two phonetically confusable versions of each base.

Word strings. The word strings consisted of words obtained from the meaningless, phonetically nonconfusable version and the meaningless, phonetically confusable version of the sentences used in the sentence repetition task. For each string, a set of five words was chosen from among the one-syllable content words of one version. (In the case of phonetically confusable versions, choice was limited to rhyming words.) Each set of five words was then rearranged to form an agrammatic sequence, a manipulation which resulted in a final set of 14 five-item agrammatic word strings, seven of which contained words that rhymed, and seven of which were made up entirely of nonrhyming words.
Procedure

Both the sentence repetition and word string repetition tasks were conducted within a single 20 minute session, with all subjects receiving the sentence repetition materials first. Transcriptions of each subject's responses were made during the experimental session by the examiner and later checked against a tape recording of the child's responses.

The test session was preceded by a training procedure designed to assure that the child understood the task. The examiner explained:

"I want you to listen to a sentence and then to try to repeat it as best you can. I'll say each sentence twice—the first time I would like you just to listen, but after the second time you hear it, you should try to repeat it. Some of the sentences may seem strange. Sometimes you may find it hard to remember all of the words. It's important for you to say as many of the words as you can remember, even if you have to guess or skip over parts you don't remember. Let's try a few, just for practice. I'll read each sentence twice. After the second time I read it, you try to repeat it. Ready?"

Following the instructions, the child was presented with a set of four practice items: two 13-word meaningful, phonetically nonconfusable sentences and two 13-word meaningless, phonetically nonconfusable ones. The experimenter read each sentence twice, after which the child was asked to repeat the sentence. If the child made no attempt to respond, the sentence was read a third time; children who hesitated over a word were encouraged to guess or to skip over that word. On completion of the four practice items, the child was advised:

"Now I am going to use the tape recorder to play some more sentences for you. This time, a man will say the sentence. He'll say each sentence twice, just as I did. Remember, try to say the sentence after you hear it the second time. Say as many words as you can."

At this time, a pre-recorded series of the test sentences was played to the child. The series included four versions of each of the seven base sentences, arranged in a fixed random order. Each sentence was repeated twice by a male, native American speaker of English, who attempted to hold prosody constant across the four versions of each base sentence. During actual testing, there was no prompting for responses, nor were unrecalled sentences repeated a third time.

The sentence repetition and word string repetition tasks were separated by a brief rest period. During this break, the examiner explained:

"Now I am going to play five words for you, one at a time. Listen to them carefully, because you will hear them only one time. After you have heard all five words, try to say them back in the same order. Remember, say as many words as you can, and guess if you have to."
The examiner then played a pre-recorded set of the 14 five-item word strings. Like the sentences, they were presented in a fixed random order, and were spoken by the same male speaker. However, unlike the sentences, each string was read only once. Words within the string were read at the rate of one per sec with prosody held neutral.

Scoring Procedure

Sentences. The error scores were the sum of omissions, substitutions and reversals made on each version of each base sentence. All versions were scored in the following manner: A score of 0 was given for correct repetition with no errors. One point was given for each word recalled in the improper sequence (relative to the preceding word), for each substitution, and for each intrusion. Words that followed substitutions or intrusions were scored relative to the immediately preceding word that had been a member of the original sentence. A score of 13 was given when a subject failed to repeat any of the words of the sentence.

Word strings. For word strings, as for sentences, the error score was the sum of omissions, substitutions and reversals. To minimize the effects of guessing, only the first five words produced during recall were counted. A score of 0 was given if all items were recalled in proper order. One point was given for each word recalled in the improper order, and for each substitution and intrusion. Words preceded by a substitution were scored relative to the immediately preceding member of the sequence. A score of five was given if the subject failed to recall any of the items.

RESULTS

This experiment was conducted to determine whether the verbal memory of good and poor readers would be differentially affected by systematic variations in phonetic confusability of the material to be recalled. For this purpose, memory for sentences and for agrammatic word strings was examined separately. The effects of systematic variations in meaningfulness were also examined in the case of sentence memory, as was the effect of variations in syntactic structure.

Sentence Repetition

In considering sentence memory, we needed first to ascertain that our good and poor readers could be differentiated by their overall performance on our materials. To this end, error score data on all the sentence versions were subjected to analysis of covariance with IQ controlled. It was found, as expected, that good readers made fewer errors overall than poor readers. The mean error score for good readers was 4.7, as compared to 5.3 for poor readers, $F_{IQ}(1,27)=7.6$, p<.01. Another prior requirement was to determine whether the stimulus variations we had introduced had any differential effects on the performances of the two groups. Each of the sentences had been presented in four versions which varied orthogonally in phonetic confusability and meaningfulness. As can be seen in Figure 1, in which mean error scores on each version type are separately plotted, children made more errors on rhyming versions than on nonrhyming versions, $F_{IQ}(1,28)=124.5$, p<.001. Meaningfulness
also produced a significant effect, \( F_{IQ}(1,28)=172.6, p<.001 \).

Having established that the good and poor readers did indeed differ in sentence memory, and that they were both affected by the stimulus variations, we turned next to the central focus of this study, which is the interaction between these stimulus variables and reading ability. We found that good readers were affected by phonetic confusability to a markedly greater extent than poor readers, \( F_{IQ}(1,27)=90.9, p<.001 \). No such interaction was obtained for the variable of meaningfulness. Supplementary analysis by t-test permitted us to further assess mean differences between the two groups on each of the four sentence versions. Here, as in the overall analysis, good readers made significantly fewer errors than poor readers on the phonetically non-confusable versions, both meaningful \( t(12)=4.2, p<.001 \); and meaningless \( t(12)=5.1, p<.005 \). In contrast, when the items were phonetically confusable, the performance of the good readers actually dropped to the level of that of the poor readers. Thus, in recall of the rhyming versions, both meaningful and meaningless, the performance of the good and poor readers did not differ significantly, as is depicted in Figure 1.

An analysis was conducted to examine the consistency of the effects of phonetic confusability and meaningfulness across the seven base sentences. In order to ascertain whether some few sentences contributed disproportionately to the main effects revealed by the analysis of covariance, we compared performance among the four versions of each base. More errors were made on rhyming versions than on nonrhyming ones for six of the seven base sentences, and more errors were made on meaningless versions than on meaningful ones for all seven base sentences. Analysis of variance reveals a significant interaction of phonetic confusability and type of base sentence, \( F(1,168)=5.9, p<.001 \), and a significant interaction of meaningfulness and type of base sentence, \( F(1,168)=8.2, p<.001 \). However, there is no three-way interaction of reading ability, phonetic confusability and type of base sentence, or of reading ability, meaningfulness and type of base sentence.

An additional analysis was carried out to treat base sentence as a random variable nested within phonetic confusability and meaningfulness (see Clark, 1973). A significant interaction of reading ability and phonetic confusability was upheld, \( \min F'_{IQ}(1,31)=4.3, p<.05 \); but there was no significant interaction of reading ability and meaningfulness.

We turned finally to compare performance across the seven base sentences, a comparison which is not central to our purposes, but is nevertheless permitted by our design. Since the base sentences were chosen to vary along a number of syntactic dimensions, it was expected that error rates in recalling them would differ. This expectation was confirmed, \( F(1,168)=29.3, p<.001 \). There was, however, no significant interaction of reading ability and the effects of base sentence, showing that good and poor readers in our sample were comparably affected by the syntactic variations. A comparison of the distribution of errors made by good and poor readers on each of the four versions of each sentence provides further evidence that the two groups reacted similarly to variations in syntactic structure. The frequency of errors as a function of the position of words in the sentence was significantly correlated for the two groups in most versions, \( r(13)>.46, p<.05 \) for 26 of
Figure 1. Good and poor readers' mean performance on meaningful and meaningless sentence versions, in nonrhyming and rhyming conditions.

Figure 2. Good and poor readers' mean performance on word strings in nonrhyming and rhyming conditions.
the 28 versions; \( r(13) = .68, p < .005 \) for 21 of them. Thus the errors of good and poor readers are similarly distributed, differing only in frequency of occurrence.

**Word strings**

As was the case with the sentence repetition data, error scores on word string recall were subjected to an analysis of covariance with IQ controlled. Mean scores for good and poor readers are plotted in Figure 2. It may be seen that an overall difference in error score was again found for good and poor readers, \( F_{IQ}(1,27) = 4.50, p < .05 \). Also apparent once more is a significant effect of phonetic confusability, \( F(1,28) = 12.8, p < .002 \): children made more errors in recall of rhyming strings. The crucial interaction of reading ability and the effect of phonetic confusability is again strongly manifest, \( F_{IQ}(1,27) = 9.5, p < .002 \). As illustrated in Figure 2, the performance of good readers was markedly impaired by phonetic confusability while that of the poor readers was not.

A test was made of the generality of these findings by an analysis of variance with word string treated as a random variable. Here, as in the preceding analysis of covariance, the interaction of reading ability and phonetic confusability is significant: \( \text{min} F_{IQ}(1,14) = 5.71, p < .05 \).

**DISCUSSION**

As we noted in the introduction, a number of studies in the research literature report that unskilled readers tend to perform more poorly than skilled readers in short-term recall of letter strings, word strings, and sentences. In studies of letter-string recall (Liberman et al., 1977; Shankweiler & Liberman, 1976), demonstrations of the greater vulnerability of good readers to the penal effects of phonetic confusability suggest that these children place greater reliance on phonetic coding as a short-term memory strategy. Correspondingly, the demonstration that recall in poor readers was little affected by the phonetic characteristics of the items suggests that they are making ineffective use of phonetic coding in working memory. Our aim in the present study was to test the generality of this interpretation by asking whether phonetic confusability will also differentially affect good and poor readers' recall not only of alphabetic strings but also of sentences and word strings.

To this end, good and poor readers in a second grade sample were asked to repeat specially designed sentences and agrammatic word strings. Consistent with previous reports, good readers were better than poor readers when the material to be recalled, whether sentences or word strings, contained no phonetically confusable words. In contrast, the performance of good readers fell to the level of poor readers when phonetically confusable words were present. Although some studies have found that poor readers are more adversely affected than good readers by manipulations that destroy meaningfulness (Pike, Note 1, Note 2; Wiig & Roach, 1975), our systematic variations of meaningfulness and syntactic structure did not differentially affect the two reading groups. The primary distinction, once again, was that good readers were severely impaired by the introduction of phonetic confusability and the
poor readers were not.

These findings confirm the results of Liberman et al. (1977) and extend them to the more natural task of sentence recall. Since the same pattern of interaction with phonetic confusability has been found for three different classes of items—letters, words and sentences—a common etiology is implicated. We follow Liberman et al. (1977) in suggesting that the poor readers' substandard recall of verbal material may be caused by failure to make effective use of phonetic coding in working memory.

We have viewed these and other findings of correlation between effective use of phonetic coding and success at learning to read as further indication of the ubiquitous involvement of speech coding in the reading process. It could be supposed, however, that ineffective phonetic coding is a by-product rather than a determinant of reading difficulty. This question might be laid to rest if it could be shown that deficient use of phonetic coding in preschool children is predictive of reading failure, both in English and in languages that manifest quite different morphologies and writing systems. We are in the process of gathering data pertinent to this issue.

Other investigators have commented on the association between reading difficulty and deficient verbal short-term memory (see, for example, Perfetti & Goldman, 1976; Perfetti & Lesgold, 1979; Vellutino, Steger, DeSetto, & Phillips, 1975; Vellutino, Steger, Kaman, & DeSetto, 1975). Moreover, we are not alone in supposing that these deficiencies apply to perception of language by ear as well as by eye. Our supposition that a number of memory related problems may be seen as manifestations of deficient phonetic coding (Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979) is consistent with the views of Perfetti and his colleagues. It is appropriate at this point to consider what precisely might be the basis of the poor reader's limitations in use of phonetic representation. In a recent paper (Shankweiler et al., 1979), we raise the question whether the deficits may extend beyond the memorial aspects of language, involving perhaps the level of perceptual encoding. If so, then sufficiently stringent tests of speech perception might be expected to distinguish good and poor readers of the sort studied here. We are currently investigating this possibility, bearing in mind the hypothesis of Perfetti and Lesgold (1979) that the short-term memory differences between good and poor readers may largely derive from slower encoding on the part of poor readers.

At the same time that we are led to consider the basis of the poor reader's ineffective use of phonetic coding, we are also led to speculate as to its broader implications. Here we are guided by the assumption that a major function of phonetic coding in both written and spoken language is to facilitate interpretation of stretches of discourse longer than the word. If poor readers do, in fact, fail to make effective use of phonetic coding, then they may have difficulty comprehending some kinds of sentences in situations in which working memory is stressed.

We conclude by suggesting two ways in which working memory might be stressed in sentence processing. First, it may be stressed when recovery of syntactic structure requires retention of many component words of a sentence.
Such could be the case in center-embedded sentences and sentences involving extensive movement or deletion (cf. Frazier & Fodor, 1978; Kimball, 1975, for a discussion of sentence parsers). Accordingly, these might pose more difficulty for poor readers than for good readers. Second, even when syntactic structure is relatively simple, working memory may be stressed if word order is in some way crucial. The importance of word order in this sense has been discussed by Baddeley (1978) and is exemplified in the Token Test of DeRenzi and Vignolo (1962). We suspect that Token Test instructions such as "Touch the large, red triangle with the small, green square," might differentiate between good and poor readers, and we intend to pursue this possibility.

REFERENCE NOTES


REFERENCES


Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer,


