Three experiments investigated the extent to which children of various intellectual abilities rely on speech-related processes in working memory. Subjects consisted of 16 learners with mental retardation (MRLs) in grades 5-9 and 16 nondisabled learners (NLs) in grades 2-3, matched in word recognition skills. The independent variable within each experiment was the phonemic content of verbal stimuli. In a word memory task, MRLs showed no decrement in performance with phonemically similar word lists, whereas NLs recalled fewer words from lists containing phonemically similar words. In a listening comprehension task, both groups evidenced similar decreases in accuracy when sentences contained phonemically similar words. In a reading comprehension task, although MRLs read more slowly overall, they showed no decrement in reading speed on sentences containing phonemically similar words, whereas NLs tended to read phonemically similar sentences more slowly than normal sentences. Results suggest that ineffective use of phonemic coding in working memory may be contributing to the reading difficulties for MRLs. (Contains approximately 60 references.) (Author/JDD)
PHONEMIC SUPPORT IN COMPREHENSION: COMPARISONS BETWEEN MILDLY RETARDED AND NONRETARDED LEARNERS

Mary Kabrich
Deborah McCutchen

University of Washington

Correspondence to:
Mary Kabrich
312 Miller Hall, DQ-12
University of Washington
Seattle, WA 98195
(206) 543-1139

This manuscript is currently under review. Please do not cite without permission.
PHONEMIC SUPPORT IN COMPREHENSION: COMPARISONS BETWEEN MILDLY RETARDED AND NONRETARDED LEARNERS

Abstract

Three experiments investigated the extent to which children of various intellectual ability rely on speech-related processes in working memory. Subjects consisted of 16 mentally retarded learners (MRLs) and 16 nonretarded learners (NLs) matched in word recognition skills. The independent variable within each experiment was the phonemic content of verbal stimuli. In a word memory task, MRLs showed no decrement in performance with phonemically similar word lists, whereas NLs recalled fewer words from lists containing phonemically similar words. In a listening comprehension task, both groups evidenced similar decreases in accuracy when sentences contained phonemically similar words. In a reading comprehension task, although MRLs read more slowly overall, they showed no decrement in reading speed on sentences containing phonemically similar words, whereas NLs tended to read phonemically similar sentences more slowly than normal sentences. Results suggest that ineffective use of phonemic coding in working memory may be contributing to the reading difficulties of MRLs.
In the last two decades, exciting progress has been made in isolating the factors that differentiate skilled and less skilled readers (e.g., Anderson, Hiebert, Scott, & Wilkinson, 1984; Perfetti, 1985). However, with few exceptions, this research has contrasted skilled readers with a single category of less able readers - those of normal intelligence who show a deficit specific to reading. A second category of less able readers are those whose reading difficulties appear in the context of low intelligence and low overall academic achievement, that is, mentally retarded readers. There is a sizable discrepancy in the attention given these two categories of less able readers within recent reading research. In an extensive literature search for the period from 1975 to 1985 Blanton, Semmel, and Rhodes (1987) found only four recent studies that directly investigated word identification skills of mentally retarded learners. In contrast, a literature search for the same period of time for the first category of readers revealed 1,336 studies.1

While a diagnosis of mental retardation customarily denotes significant and pervasive developmental delays, there exists within this group relatively broad variation in the level of reading skill acquisition. Furthermore, despite widespread reading deficiencies among mildly mentally retarded learners (MRLs), the literature provides little in the way of convincing, empirically based explanations for these phenomena (Blanton et al., 1987). The present study provides a starting point from which an empirically based explanation for the reading deficiencies of this population may be advanced. Specifically, in this paper we will develop the view that failure to actively maintain phonemic codes in memory might be but one aspect of a more general problem that affects MRLs' performance on both memory and reading tasks. We will review relevant literature that examines: 1) the link between reading ability and short-term memory performance with nonretarded children, 2) MRLs' memory limitations and the implications these may have for reading, and 3) the contribution of phonology to comprehension.

We will report 3 experiments that compare 16 MRLs with a matched control group of

1 This number was obtained using the search capacities of Psyc. Lit.
nonretarded learners (NLs). The results of these experiments expose theoretically significant
differences and similarities in phonemic processing between our two subject groups, shedding
some light on how memory processes may impact reading ability. The primary question of
interest was the extent to which MRLs exploit phonemic information in short-term memory
during language processing.

**Reading Ability and Short-Term Memory Processes in Nonretarded Learners**

One of the more consistent findings in the literature on individual differences in reading
is the relationship between reading ability and short-term memory (STM), with better reading
skill being associated with superior performance on STM tasks (Daneman & Carpenter, 1980;
Heulsman, 1970; Turner & Engle, 1989). Deficits in STM for individuals with reading problems
has been demonstrated on a number of different measures, i.e., digit span, letter strings, sentence
tasks, and recall for pictures of familiar objects (Heulsman, 1970; Perfetti, 1985; Stanovich,
1982; Vellutino & Scanlon, 1985). Interestingly, the memory problem for poor readers generally
has been found to be specific to memory tasks involving verbal material or material that is easy
to represent linguistically (Heulsman, 1970; Stanovich, 1982; Torgesen, 1988; Torgesen &
Houck, 1980).

One promising explanation of this relationship posits that poor readers utilize less
efficient phonemic coding processes than do skilled readers (Mann, Liberman, & Shankweiler,
1980; Shankweiler & Crain, 1986). Such an account has the advantage of explaining a number of
characteristics of poor readers, including difficulties in understanding spoken sentences,
decoding words, and comprehending text. Research provides evidence that all these skills to
some degree rely on verbal memory (see Shankweiler & Crain, 1986).

Support for this explanation has come largely from studies showing that poor readers are
less affected than good readers by phonemic similarity in memory tasks. These studies are
patterned after Conrad's seminal study (1964) in which subjects recalled sequences of letters, some of which rhymed. The typical paradigm employed in these studies is either a visual or aural presentation of a string of words, letters or sentences containing either rhyming or phonemically distinct phonemes. Of interest is the pattern of performance between skilled and less skilled readers on the two types of stimuli. The expectation is that good readers are better able to form phonemic representations than poor readers and will, as a result of phonemically-enhanced memory codes, recall more of the phonemically distinct items than the poor readers. However, good readers' sensitivity to the phonetic content will hinder their ability to recall those items which are not phonemically distinct. Therefore, the expectation is that the good readers will show poorer performance on the rhyming items while poor readers will be unaffected by phonetic similarity. This prediction has been confirmed by a number of studies (e.g., Mark, Shankweiler, & Liberman, 1977; Mann et al., 1980). However, more recent studies have challenged these findings, arguing that poor readers' apparent phonemic deficits are symptoms rather than causes of their limited short-term memories. Holligan and Johnston (1988) demonstrated that poor readers show phonemic similarity effects in tasks that do not overburden memory and fail to do so only in tasks that tax memory or require relatively sophisticated segmentation skills. The argument is that poor readers can, in limited circumstances, access phonemic information, but they do not routinely use it as good readers do.

**Memory and Mental Retardation**

Whereas much of the current literature on reading disabilities has focused on the relationship between memory functioning, speech processes and reading ability, exploration of these linkages in the field of mental retardation is scarce. This is not to say that these issues have been ignored by researchers. Memory functioning is an area of research that has received considerable attention in this field (e.g., Belmont & Butterfield 1977; Belmont, Ferretti &
One of the most firmly established cognitive characteristics of MRLs is a deficiency in the use of memory strategies, such as rehearsal (e.g., Detterman, 1979). However, training studies have demonstrated that in many instances it is the intention to use an appropriate strategy which is absent, not necessarily the ability to do so (Campione, Brown, & Ferrara, 1982).

Another aspect of memory functioning which has been correlated with intelligence is the set of abilities referred to as speed of processing, but this aspect of memory functioning has received far less attention than the research on strategic behaviors. In general, speed of processing refers to the speed with which various types of mental operations take place. It has typically been assessed by tasks that measure time required to search short-term memory, access overlearned long-term memory codes, or identify incoming stimuli. In the most detailed set of studies in the area, Hunt and his associates (Hunt, Frost, & Lunneborg, 1973; Hunt, Lunneborg, & Lewis, 1975; Hunt, 1978) have found that college students with high and low verbal ability differ from each other in the speed with which they carry out various mental operations. Hunt and associates found that subjects with higher verbal ability, as measured on standardized tests, were faster to judge letter identity, but especially so when the identity judgments were based on names rather than visual features. This was interpreted as indicating that faster memory access to name codes is a basic information processing skill that underlies verbal ability. In subsequent work by Hunt and his colleagues (reviewed in Hunt, 1978), these findings were extended to a sample of mildly MRLs.

Several other studies (e.g., Harris & Fleer, 1974) have found that MRLs are slower than NLs in searching STM. For example, when subjects are shown a simple probe item following a set of items they were asked to memorize, MRLs take longer to decide if the probe was from the item list than do NLs. This longer reaction time has been interpreted as an indication of slower memory scanning rates.
Implications for Mentally Retarded Readers

These findings are of particular interest in light of the relationship between encoding rate and working memory limitations in reading. Research over the past two decades has also linked speed of object naming to reading ability (e.g., Manis, 1985; Perfetti, 1985; Shankweiler & Crain, 1986). A number of researchers (Baddeley, Thomson, & Buchanan, 1975; Dempster, 1981; Huttenlocher & Burke, 1976) support the general conclusion that slower rates in item identification result in a shorter memory span since relatively less capacity is left over for storing items. Furthermore, problems with both object naming and memory capacity have been claimed as symptomatic of problems at the level of phonology (Shankweiler & Crain, 1986). While the link between verbal coding, phonology and reading has not been directly investigated with MRLs, it appears reasonable to likewise consider MRLs' difficulties with object naming, memory, and reading as potential manifestations of failure to use phonetic coding effectively.

Thus, there have been numerous investigations of MRLs' memory processes, but most studies have concentrated on specific strategies (e.g., rehearsal) and their acquisition rather than the relation of memory processes to reading ability. In addition, few studies have investigated word identification skills among MRLs. The majority of those have done so in the context of evaluating training techniques, without directly exploring cognitive processes in relation to word recognition skills (e.g., Ault, Gast, & Wolery, 1988; Milgram & Riedel, 1969; Neville & Vandeveer, 1973; Singh & Singh, 1988).

One notable exception is a study by Mason (1978). Mason sought to answer the question of whether or not MRLs use reading strategies similar to those used by NLs. Mason's findings suggested that MRL's use a strategy of memorizing whole words, whereas nonretarded subjects used letter-sound information to pronounce words.

Mason's results are only partially supported by the findings of McCutchen, Dibble, Kabrich, and Laird (1989) who compared MRLs reading at first and second grade levels with
nonretarded first and second grade students on several phonemic awareness and word reading tasks. Despite significant overall differences in reading speed, which favored the nonretarded students, both the MRLs and NLs were faster when reading words containing regular letter-sound correspondences than those containing irregular correspondences. Theoretically, it is use of letter-sound information that results in faster reading times for regularly spelled words; and since both groups showed a regular-word advantage, both must use letter-sound information during reading. However, McCutchen et al. (1989) found different patterns of performance between groups when they asked students to analyze words at the sub-syllabic level. While both groups were equally able to delete a syllable segment (i.e., saying baseball with out the base), the MRLs were less accurate (95% vs. 70%) than the nonretarded students when the deleted segment was a single phoneme (i.e., saying meat without the /m/). MRLs evidenced a similar disadvantage compared to NLs when asked to judge whether words shared the same sound in either initial, medial or final position. While McCutchen et al. (1989) did not see evidence that MRLs were totally insensitive to letter-sound information, their results, taken together with those of Mason (1978), suggest that MRLs may not use sub-syllabic phonemic information to the same extent as NLs.

From Word Reading to Comprehension

A number of the earlier findings in the literature have reported that MRLs read commensurate with their mental-age grade expectancy at the primary reading level (Bennett, 1932; Blake, Aaron, & Westbrook, 1969). However, this does not continue to be the case at the intermediate reading level. At the intermediate reading level (third to fifth grade), MRLs achieve below their mental age expectancy and below reading levels of NLs of similar mental ages (Blake et al., 1969; Bliesmer, 1954; Sheperd, 1967). Although this decline has been attributed to difficulties in comprehension (Blake et al., 1969; Bliesmer, 1954; Dunn, 1956), the nature of
these difficulties has not been clearly specified. The research emphasis in general has been on a variety of instructional strategies aimed at improving reading comprehension skills (Blanton et al., 1987). While these studies may indeed prove helpful to MRLs, in order to gain a more complete assessment of MRLs reading difficulties, it will be necessary to obtain a more precise understanding of the processes that support reading comprehension.

In a recent study, Merrill and Bilsky (1990) investigated the ability of MRLs to construct semantic representations of sentences. Their results indicated that MRLs are deficient in their ability to construct integrated representations of single sentences. An important question left unanswered by this study is the cognitive processing abilities that enable some individuals, relative to others, to construct better integrated sentence representations. Merrill and Bilsky (1990) suggest that one factor to consider is the speed with which verbal coding processes can be executed. A second possible factor is the nature of the verbal code itself. While Mason's (1978) study suggested that MRLs depend on whole-word retrieval as a word identification strategy, there still remains the question of whether phonemic information plays a role in accessing whole words and supporting subsequent comprehension processes.

The work of McCutchen and associates (McCutchen & Perfetti, 1982; McCutchen, Bell, France, & Perfetti, 1991) has provided evidence that phonemic information is involved in silent reading. Perfetti and McCutchen (1982) proposed a model of reading whereby phonemic information plays a critical role in the comprehension of sentences. They argued that phonemic information is part of the code held in memory at least briefly following lexical access. This code provides an index for referencing lexical items and for maintaining order information. (See also Shankweiler & Crain, 1986.) Indirect evidence for this model has been provided by experiments employing tongue-twisters in sentence reading tasks. The rationale guiding these experiments is similar to that of the memory tasks described previously: If phonemic information is used as a code index to lexical items in memory, the code will become vulnerable to
interference when phonemic distinctiveness is lost as a result of phonemic repetition. Evidence of such interference has come from studies demonstrating longer reading rates for sentences containing tongue-twisters (McCutchen & Perfetti, 1982; McCutchen et al., 1991). McCutchen and colleagues argued that the ability to represent the sounds of words in memory facilitates normal comprehension (i.e., comprehension of text containing a normal mix of phonemes).

There have also been several attempts to use the tongue-twister paradigm to examine individual differences (Crain-Thoreson, 1991; Crain-Thoreson & McCutchen, 1989). Surprisingly, given the findings of Mann et al. (1980), these studies failed to find differences in the magnitude of the tongue-twister effect evidenced by readers of different ability. Less skilled readers were as sensitive as skilled readers to phoneme repetition while reading sentences.

Whereas the literature on NLs has focussed on the roles memory and speech processes play in reading, reading research with MRLs has not paralleled that emphasis. In fact, studies exploring the relationship of phonemic coding and reading skill with this population are almost nonexistent. In this study, we investigate the relationship between memory, phonemic processes and comprehension with a sample of MRLs.

We begin by assuming that the processing of linguistic information proceeds along a continuum of speech activation, as argued by Perfetti and McCutchen (1982). Their activation model is reproduced in Figure 1. The partial phonemic representations that are activated routinely during silent reading may be represented at the lowest level of activation on such a continuum. As suggested by the studies of Lukatela and Turvey (1990), as well as many others, this level can most aptly be described as automatic and, as such, may be outside of strategic control (see McCutchen & Perfetti, 1982). Somewhat higher on the continuum are more complete phonemic representations (sometimes referred to as vocalic or acoustic images, see Baddeley & Hitch, 1974), which are less automatic and more amenable to strategic interventions,
such as rehearsal. At higher levels of the continuum are actual articulatory motor programs for
subvocal and finally overt speech.

Insert Figure 1 about here.

In order to address the extent to which MRLs exploit phonemic information in STM, we
compared their performance to that of a matched control group on three separate but related
tasks that evoked different levels of phonemic processing. The first task (Experiment 1) required
an intermediate level of activation on the continuum just described: subvocal rehearsal of a
random list of words for verbatim recall. This memory task replicated and extended the work of
Mann et al (1980). The second task (Experiment 2) was a listening comprehension task. Here
subjects listened to sentences that varied in phonemic content and then made acceptability
judgments. This task did not demand explicit rehearsal because verbatim recall was not
necessary. In the third task (Experiment 3), subjects read and judged sentences in which
phonemic information was systematically varied. Here, a more abstract code -- near the low end
of the continuum of activation -- needed to be activated and maintained by subjects in order to
read and comprehend sentences.

EXPERIMENT 1

Experiment 1 involved a simple word memory task in which the nature of the phonemic
content in the words was varied. This task was adapted from the word memory task of Mann et
al. (1980) in which skilled and less skilled readers were compared on their recall of
phonemically controlled word strings. A major criticism of that study (e.g. Holligan & Johnston,
1988; Johnston, 1982; Johnston, Rugg, & Scott, 1987) was that task difficulty was not equal for
both groups of readers. Because Mann et al. matched readers on chronological age instead of
reading level, a disparity in task demands existed for the two reader groups. In Experiment 1, we matched subjects in terms of word recognition skills, although no explicit matching was done in terms of STM capacity (see Holligan & Johnston, 1988). If retarded and nonretarded learners, matched in reading level, differ in their utilization of phonemic encoding in working memory, we are likely to observe differences in the sensitivity of the two groups to phonemically confusing word lists.

Experiment 1 also differs from the Mann et al. (1980) study in the addition of a second phonemic category -- alliteratives. Our previous study (McCutchen et al., 1989) indicated that on a phonemic awareness task, both second grade NLs and MRLs from grades 4-6 were more accurate on judgments concerning beginning sounds than those concerning rhymes. By including both alliteratives and rhymes, we may be in a better position to take a fine-grained look at the various aspects of phonemic codes used in memory.

Of the three experiments comprising the present study, Experiment 1 is most subject to rehearsal affects. Given the vast literature on MRLs' tendency not to engage in spontaneous rehearsal (see Campione & Brown, 1977), we hypothesized that the retarded readers would not be adversely affected by the phonemic content of the item sets. However, for the NLs, we expected the pattern to be different. We expected that they would actively engage in strategies to keep the verbal code active in STM. This greater reliance on phonemic representation as a means of remembering words should result in greater vulnerability of the NLs to the effects of phonetic confusability. Furthermore, since rhymes contain two repeated sounds, rather than just one as in the alliteratives, we expected that NLs would be more adversely affected by the rhymes than the alliteratives.
Method

Subjects

The subjects consisted of 16 mildly mentally retarded learners, grades 5 - 9 and 16 nonretarded 2nd and 3rd grade children selected from schools in a large urban district. The MRL sample had a mean chronological (CA) age of 160.5 months and a mean IQ of 66. All IQs were based on recent administrations of the Wechsler Intelligence Scale Revised for Children. Although no IQ information was available for the NLs, their mental ages were estimated to be roughly equivalent to their CAs. No children receiving special education or remedial reading services were included in the nonretarded group.

The first stage in selecting the mildly mentally retarded learners consisted of identifying children whose school records indicated IQs of 75 and below, as well as some indication of at least second grade word recognition skills. Once identified, these children were then given The Peabody Individual Achievement Test (PIAT) for Word Recognition (Dunn & Markwardt, 1970), an untimed test that requires subjects to read aloud words increasing in difficulty. (The test-retest reliability of the PIAT is reported to be .89.) NLs were then matched on the basis of PIAT raw reading scores. The mean average raw score on the PIAT for both the MRLs and the NLs was 50 (3.0 grade level score). The standard deviation was 8.76 for the NLs and 9.55 for the MRLs. The reading grade level scores for both groups ranged from 2.0 to 5.0.

Materials

The materials for Experiment 1 consisted of 21 sets of 5 word strings. The sets contained words of three types: rhymes, alliteratives, or a natural mix of phonemes (control set). The words used were one-syllable content words matched in frequency (Carroll, Davies & Richman, 1971). Fourteen of the sets were composed of words from the control and rhyme sentences used in the Mann et al. (1980) study. The final seven word sets were composed of alliteratives. All word
strings were arranged in an ungrammatical sequence. The sets of words were prerecorded in 2 fixed random orders -- one the inverse of the other -- at a rate of approximately one word per second. Half of each subject sample listened to each order. Table 1 provides example word strings from each phonemic category.

---

Insert Table 1 about here.

---

Procedure

Prior to the presentation of the word strings, the examiner explained to the child that he/she was to listen to a group of words and then try to repeat as many as possible. The child was told that it was OK to guess or skip words. The child was then presented with a set of 2 practice items, followed by the test items. Each word set was presented once. Between the presentation of each set, the examiner paused the tape recorder to ask if the subject was ready. The subjects' responses were recorded both on tape and on paper by the examiner. Subjects' responses were scored on the basis of the total number of correctly recalled words, regardless of order, in each of the three word categories. Thus, a perfect score for any one category (controls, alliteratives, or rhymes) would be 35 (5 words per set X 7 sets). Subjects were tested individually, with each session lasting approximately twenty minutes.

Results and Discussion

The number of correctly recalled words was analyzed using a multivariate repeated measures analysis of variance (ANOVA). The data were analyzed in a 2 (subject group) X 3 (word category) factorial design, with word category (control, alliterative, rhyme) as the within subject measure, and subject group (retarded, nonretarded) as the between subject measure. The mean scores for the word categories for each subject group are presented in Figure 2.
The ANOVA showed no main effect of group, with the NLs and MRLs correctly recalling approximately the same number of words (23 compared with 21, respectively). The main effect for sentence type was significant (F(2,29)=12.98, p<.001); however, this main effect was compromised by a significant interaction (F(2,29)=7.01, p=.003), reflecting the fact that the NLs were much more affected by phonemic similarity than were the MRLs (especially by the rhyme manipulation). Whereas the NLs correctly recalled 25.5 control words, they recalled 23.4 alliterative words and only 9.5 of the rhyme words. The MRLs however recalled 21.2 control words, 21.2 alliterative words and 20.3 rhyme words.

These results provide support for the hypothesis that the MRLs are less affected by the phonemic content of the words than the NLs. The results also support the hypothesis that NLs are more affected by rhymes than by alliteratives. While there is some suggestion in Figure 2 that alliteratives also affect memory for the NLs, only the difference between rhymes and controls reached significance (p < .05 in follow-up tests).

As Mann et al. (1980) found with nonretarded poor readers, MRLs in the present study were not affected by the phonemic content of words on a simple memory task. While the interpretation offered by Mann et al. has been criticized on the basis of task difficulty (Holligan & Johnston, 1988; Johnston, Rugg, & Scott, 1987), an explanation based on task difficulty seems unlikely here. With the sample reported here, no significant difference in the overall level of accuracy was found between the two groups. Furthermore, both groups' rate of accuracy exceeded 60%, in contrast to the low overall accuracy levels (roughly 30% to 40%) of poor readers reported in previous studies showing reduced phonemic effects (Shankweiler, Liberman, Mark, Fowler, & Fischer, 1979; Siegal & Linder, 1984).
A more plausible explanation for the interaction observed in this experiment is that the MRLs, unlike the NLs, did not engage in rehearsal. Although this explanation can be readily supported in the memory literature on MRLs (e.g., see Detterman, 1979), it may not be the only explanation, nor the best explanation (considering the relatively high accuracy of the MRLs). Nonretarded poor readers who fail to show phonemic sensitivity have been found to actively engage in rehearsal strategies (Mark, Shankweiler, Liberman, 1977). Therefore, while lack of rehearsal may be contributing to these findings, the nature of the activated word codes may also be different between MRLs and NLs. Perhaps, as Mason (1978) suggested, MRLs make less use of phonemic properties of words. The next two experiments take a closer look at the phonemic codes used by MRLs and NLs in tasks that do not necessarily entail explicit rehearsal.

EXPERIMENT 2

The purpose of Experiment 2 was to investigate phonemic processing at a level more subtle than that required in the memory task in Experiment 1. In Experiment 2, the subjects listened to sentences that varied in phonemic content and made acceptability judgments based on the sentences' meaningfulness. Thus, this was a listening comprehension task. While listening comprehension does not require explicit rehearsal, words do need to be remembered long enough to construct meaning. In Experiment 2, words needed to be remembered long enough to judge whether the sentence as a whole was meaningful.

The hypotheses for Experiment 2 were similar to that of Experiment 1, at least for the nonretarded learners. Once again, we expected that the NLs would rely on phonemically enhanced lexical representations in memory and this reliance would cause them to be vulnerable to the effects of phonemic confusability. Furthermore, because vowels have longer acoustic duration and contain two repeated sounds, we expected that vowels would be more likely than alliteratives to interfere with comprehension. Hypotheses regarding MRLs were less clear. If the
results of Experiment 1 were due to MRLs failing to rehearse, then differences between groups might not occur in Experiment 2, because sentence comprehension does not require explicit rehearsal. However, if MRLs fail to represent phonemic information as part of the word codes used in comprehension, then they might again show no phonemic confusability effect.

Subjects
The subjects were the same as those in Experiment 1.

Materials
Items for the sentence acceptability task consisted of 21 semantically acceptable (meaningful) sentences and 15 semantically unacceptable sentences. The 21 semantically acceptable sentences were composed of 7 sentence triads, with all members of a triad containing a similar syntactic and semantic structure but each varying in phonemic content (rhymes, alliteratives, and controls). Two of the three members of each triad came from the set of "meaningful, phonetically nonconfusable" sentences (the controls) or the set of "meaningful, phonetically confusable" sentences (the rhymes) used in the Mann et al. (1980) study. The third member of each triad consisted of a sentence containing alliteratives.

Each version of a given base structure (rhyme, alliterative, and control) was constructed by substituting content words, with function words held constant. In addition, all versions of each base sentence were matched with respect to word frequency (Carroll, Davies & Richman, 1971) and syllable length. The semantically unacceptable sentences were constructed by switching the position of two words within an acceptable sentence to render the sentence grammatically unacceptable. In order to reduce the amount of time for this task, only 5 of the 7 meaningful sentences in each set were turned into unacceptable items. All three versions of the unacceptable sentences were matched with respect to the position and function of the words
chosen to be switched. This resulted in a maximum score of 24 (7 acceptable + 5 unacceptable x 2 repetitions; see Procedure) for control, rhyme, and alliterative sentences. Table 2 presents an example of one of the base forms presented in the three different versions of semantically acceptable and unacceptable sentences.

---

Insert Table 2 about here.

---

The sentence sets were prerecorded in two random orders, one the inverse of the other, at a rate of approximately one word per second. Half of the subjects in each group were randomly assigned to each order.

**Procedure**

Subjects listened to sentences presented on a tape recorder and decided whether or not each sentence made sense. Prior to the presentation of the sentences, the subjects were given 4 practice trials with feedback to ensure that they understood the task. The subjects were told that the person reading the sentences into the tape recorder sometimes mixed two words around so that the sentence did not make sense. In order to increase reliability, each sentence was presented two times in sequence. Subjects were told that they would have a second opportunity to hear each sentence, however they should still respond upon first hearing the sentence in one of three ways: "Yes" if it makes sense; "No" if it doesn't make sense; or "I don't know." Subjects were told they could change their answer if they wished, after hearing the sentence a second time. Subjects were also cautioned to listen to the entire sentence before making a judgment because some sentences might sound similar.

During the actual testing there was no prompting for responses. After recording the child's response to each sentence, the examiner stated either "Now listen again" or "Here is a
new one" before presenting the next sentence. Accuracy scores were calculated by summing the total number of correct responses to all sentences on both trials. Only when the child clearly indicated that a sentence either made sense or did not make sense were points awarded. Thus, if upon hearing the sentence a first time, the child responded "I'm not sure" or "I didn't hear it," and then on the second trial stated "yes" to a correctly constructed sentence, they would receive 1 point was awarded for the "yes" response. Subjects were tested individually, with each session lasting approximately twenty minutes.

Results and Discussion

The number of correctly identified sentences were analyzed using a multivariate repeated measures analysis of variance (ANOVA). The data were analyzed in a 2 (subject group) X 3 (sentence category) factorial design with sentence category as the within-subject measure, and subject group as the between-subject measure. The mean score for the sentence categories for each subject group are presented in Figure 3.

The analysis revealed an overall main effect for sentence type (F(2,29)=28.30, p<.001). Follow-up tests indicated that control sentences for both subject groups differed significantly from rhymes (F(1,30)=54.54, p<.001) and alliteratives (F(1,30)=10.73, p < .01).

There was no main effect for group, with both groups correctly judging a similar mean number of sentences (15 for MRLs compared with 17 for nonretarded subjects), nor was there a significant interaction between group and sentence type. As Figure 3 illustrates, both groups of subjects were less accurate judging rhyming and alliterative sentences than control sentences.
Both retarded and nonretarded subjects were penalized by phonemic repetition in the sentences, and to similar degrees.

These results lend support to the hypothesis that both groups of subjects experience some level of phonemic activation. Unlike the word memory task in which subjects were asked to remember the words, this comprehension task did not demand explicit rehearsal. Results from Experiment 2 suggest that when MRLs are given the complete phonemic representation of words (via aural presentation), they show phonemic confusions in a comprehension task. Thus, it seems that MRLs' insensitivity to the phonemic content in Experiment 1 was not due to their complete inability to use phonemic codes.

Because the aural presentation of the sentences provided all subjects with complete phonemic representations of all the words, Experiment 2 does not yield much information as to the nature of the phonemic codes that MRLs and NLs activate on their own during reading. However, these data do yield information regarding the ability of the two groups to maintain phonemic information in memory once there, in the absence of strategic intervention. When asked to comprehend sentences, rather than recall verbatim word lists, NLs and MRLs showed evidence of similar maintenance of phonemic information. That is, the fact that the two groups showed similar accuracy rates and similar effects of phonemic confusability suggests that, once activated, the basic decay function of phonemic codes is similar for both groups. NLs may be more likely than MRLs to use rehearsal or other strategies to postpone decay of phonemic information (as suggested by the results of the memory task in Experiment 1), but in the absence of such strategic intervention, phonemic information seems equally available to both groups.

The fact that both groups of readers were more affected by the sentences containing rhymes than those repeating initial consonants is of interest. One explanation for this difference lies in the acoustic duration of vowels and the nature of rhymes. With the aural presentation of sentences in Experiment 2, the acoustic features of the stimuli were especially salient, and, as in
Experiment 1 (and in Crain-Thoreson, 1991), rhyming stimuli produced larger effects than alliterative stimuli. What is most noteworthy in this task, however, is that both the rhyming and alliterative sentences induced more errors in both subject groups, suggesting that phonemic codes play a role in comprehension for retarded as well as nonretarded learners, at least in a listening comprehension task.

EXPERIMENT 3

Experiment 3 examined whether the pattern of results observed in Experiment 2 generalized to reading as well as listening comprehension. Theoretically, when reading, it may be possible to ignore the phonemic properties of sentences, whereas this is not possible when listening. Such a view is compatible with a theoretical view of reading that involves only visual information in the access of lexical representations. While there remains considerable debate over the role of phonemic information in lexical access (e.g., van Orden, 1987), mounting evidence suggests that nonretarded adults (e.g., McCutchen et al., 1991) and children (Crain-Thoreson, 1991; Crain-Thoreson & McCutchen, 1989) use phonemic information during comprehension of sentences in silent reading. The question of interest in Experiment 3 was whether this would be equally true of MRLs. When asked to read and comprehend, and thus activate and maintain phonemic codes on their own, would MRLs show the same pattern of performance as NLs (as they did in the listening task of Experiment 2) or would they show a markedly different pattern (as they did in the memory task of Experiment 1)?

Method

Subjects
The subjects were the same as those in Experiments 1 and 2.
Materials

Items for the sentence acceptability task consisted of 24 control sentences and 24 experimental sentences adapted from materials used by Crain-Thoreson (1991). In each of the 24 control and experimental sentences, 16 were semantically acceptable and 8 were semantically unacceptable.

The 24 experimental sentences were of two varieties. Half of these sentences had content words that rhymed, and the remaining half had content words that began with the same phoneme (alliteratives). The 24 control sentences were also of two varieties. While all control sentences contained a natural mix of phonemes, half of these sentences were semantically and syntactically matched to the experimental rhyme sentences and the other half matched semantically and syntactically to the experimental alliterative sentences.

The matched experimental and control sentences were constructed with equivalent word frequencies. Relatively high frequency words in third-grade reading material (the lowest norms provided in Carroll, Davies & Richman, 1971) were used as content words, with second-grade basal readers as a guide. To create the unacceptable sentences, words from acceptable sentences were shuffled to form ungrammatical arrangements. Table 3 presents an example of the two varieties of experimental and control sentences in both the acceptable and unacceptable versions.

------------------------
Insert Table 3 about here.
------------------------

Procedure

Subjects read sentences presented in random order on a microcomputer and decided whether or not each sentence made sense. Prior to the presentation of the sentences, subjects were given 6 practice trials with feedback to ensure that they understood the task. Subjects were
told their task was to teach "Howie the computer" to write good sentences. To do this, they were told to decide whether each sentence printed on the screen was "wrong" (i.e., unacceptable) or "OK." Subjects were told to read the sentences silently and then indicate their decision by pressing the appropriately labeled button on the computer keyboard.

Sentences stayed on the screen until subjects made a response, and after each response, the next sentence appeared in 750 milliseconds. Response times and accuracy data were recorded. Subjects were tested individually, with each session lasting approximately twenty minutes.

Results and Discussion

Reading times and error rates were each analyzed using multivariate repeated measures analysis of variance (ANOVA). The data were analyzed in a 2 (subject group) X 4 (sentence category) factorial design with sentence category (rhyme control, rhyme experimental, alliterative control, and alliterative experimental) as the within measure, and subject group as the between.

Reading times. For the sake of ecological validity, only semantically acceptable reading times for sentences were analyzed. Previous experiments using the acceptability judgment paradigm have suggested that semantically unacceptable sentences may involve processes beyond simple reading and verification (McCutchen et al., 1991).

The analysis revealed an overall main effect for subject group [F(1,30)=5.94, p<.03]. The NLs read sentences more quickly than the MRLs (6.93 seconds compared with 8.55 seconds). This difference is especially interesting considering that the groups were matched on single-word reading accuracy (however, reading speed was not considered in the PIAT). There was no overall main effect for sentence type. However, a marginally significant interaction (F(1,30)=3.66, p=.065) indicated that the MRLs and NLs responded differently to phonemic similarity. Figure 4 depicts the response patterns across sentence types for the two groups. The
NLs replicated the response patterns reported in previous studies (Crain-Thoreson & McCutchen, 1989; Crain-Thoreson, 1991; McCutchen et al., 1991), showing slightly longer response times to the phonemically confusing sentences. The MRLs, however, showed no difference between alliteratives and controls, and they actually read rhymes more quickly than controls.

These results are difficult to interpret conclusively, in light of the marginal nature of the interaction. Still, Experiment 3 provides suggestive evidence that MRLs and NLs are differentially affected by the phonemic content of words when reading. Furthermore, a comparison of Figures 2 and 3 suggests that the performance picture for the two groups differs markedly across reading and listening. When reading, the two groups look very different (Figure 4), but when listening they look remarkably similar (Figure 3).

Accuracy. An initial analysis of accuracy of sentence judgment included only acceptable sentences so as to allow comparisons with the reading time data. A 2 (subject group) X 4 (sentence type) repeated-measures ANOVA was performed on accuracy scores, with subject group (nonretarded, retarded) as a between-subject factor and sentence type (rhyme, rhyme-control, alliterative, alliterative-control) as a within-subject factor. The analysis revealed no significant main effect for group or sentence type, and no interaction. In general, error rates were similar across both groups and all sentence types, ranging from 9% to 20%.

Informal inspection of performance on the unacceptable sentences suggested a somewhat different pattern of performance, so a second analysis was performed on accuracy data for acceptable and unacceptable sentences combined. The analysis of the combined data revealed no
effect of sentence type but a marginally significant effect of group (F(1,30)=3.21, p=.08), with NLs making slightly fewer errors than MRLs. Again there was no interaction.

A comparison of the two analyses provides interesting insights into the performance of the two groups. When data from the unacceptable sentences are included in the analysis, accuracy rates of nonretarded subjects are consistently higher, compared to their accuracy on acceptable sentences alone, but accuracy rates of retarded subjects are consistently lower. The differences are admittedly small, and they lead to only a marginally significant difference between subject groups. Still, the different trends suggest that the two groups adopted somewhat different response strategies. The NLs may have tended to respond "no" when in doubt (thus yielding higher overall accuracy when unacceptable sentences were included), whereas the MRLs may have tended to respond "yes" (thus yielding higher overall accuracy when only acceptable sentences were analyzed).

It is also interesting to note that the MRLs were least accurate on the rhyme sentences, the very sentences that they responded to most quickly. Thus, in addition to different response biases between the two groups, there also seems to be a speed-accuracy trade-off that may explain the MRLs' rapid response times for rhyme sentences.

All in all the results of Experiment 3 are far from conclusive. The lack of significant phonemic similarity effect in response times for the NLs is at odds with other findings (e.g., Crain-Thoreson & McCutchen, 1989; McCutchen et al., 1991). The nonretarded subjects in Experiment 3 did show tendencies to read the rhymes and alliterative sentences more slowly than their respective controls (see Figure 4), and the lack of significance may be the result of a lack of power. Only 16 nonretarded subjects participated, and the number of sentences was kept low to minimize subject fatigue. The addition of a second type of phonemic confusion (rhyming sentences) further reduced the number of items contributing to each cell mean. These and other
factors make the interpretation of a phonemic similarity effect for the NLs (or lack thereof) less than clear.

Despite potential power issues, Experiment 3 did yield a significant main effect of group, with NLs judging the sentences more quickly than MRLs. The faster reading rates of the NLs, together with their tendency toward higher accuracy (when "No" responses are included), are especially noteworthy in light of the fact that the groups were matched on word-reading accuracy. The comprehension demands of this sentence reading task produced differences between the groups that were not apparent when they read single words on the untimed PIAT, nor when they listened to sentences in Experiment 2. How might we account for these differences? One possible explanation rests on the argument that reading ability is best assessed by measures of both accuracy and speed (e.g., Perfetti, 1985). According to this explanation, our untimed measure of word reading accuracy (the PIAT) did not permit us to match our two groups adequately in terms of reading ability. While there remains some chance that the matching was not perfect, similar patterns of overall accuracy between groups on all three tasks suggests that the matching was not wildly inaccurate.

A second, more theoretically interesting explanation involves differences between reading individual words and comprehending sentences. What reading comprehension demands, in contrast to reading single words or listening to sentences, is both activating abstract phonemic codes and maintaining them until comprehension processes no longer require verbatim information (i.e., the point at which the relevant propositions have been built, as described by Kintsch & van Dijk, 1978). It may be the demand to both activate and maintain abstract phonemic codes that is the source of difficulty for MRLs.

Unlike single-word reading, sentence comprehension requires that multiple lexical representations be kept in memory simultaneously, at least until gist (propositional, or otherwise) can be derived. And in contrast to listening comprehension, reading comprehension involves
lexical representations accompanied by more abstract phonemic information, rather than the more complete phonemic representation that is provided by the acoustic stream during listening. Maintenance may be more critical in reading comprehension because abstract phonemic codes, being by their very nature less fully activated than acoustic representations, will decay sooner and thereby become unavailable to support memory and comprehension. This theoretical framework provides a context in which to interpret all three experiments described here.

**GENERAL DISCUSSION**

This study investigated the extent to which MRLs and NLs, matched in reading level, rely on speech-related processes in working memory. The two groups of readers were compared on three tasks that differed in the level of phonemic activation required. Experiment 1 compared the two groups on an explicit memory task. In such a task, rehearsal is an effective strategy, one that NLs seemed to employ more than MRLs. This inference is based on the observation that the NLs were more affected than the MRLs by phonemic similarity in the stimulus materials. From these data, however, it is not clear whether MRLs cannot activate phonemic codes or simply do not activate such codes as part of a rehearsal strategy. MRLs' failure to rehearse is well documented in the literature. Failure to maintain the phonemic codes in memory might be but one aspect of a more general problem that affects MRLs' performance on both memory and reading tasks. This was explored in Experiments 2 and 3.

Neither Experiment 2 nor 3 required explicit rehearsal, but both required memory to the extent that words must be retained, at least temporarily, as part of comprehension. Unlike Experiment 1, Experiment 2 showed similar patterns of performance between the two subject groups. When listening to sentences, both retarded and NLs were adversely affected by phoneme repetition, in rhymes as well as in alliterative sentences.
Given the pervasive differences between MRLs and NLs documented in the literature, the lack of difference between the two groups on a rather sophisticated sentence comprehension task is worth close examination, especially since a subset of these stimuli (the rhymes and controls) had produced markedly different patterns of recall in good and poor readers of average intelligence (Mann et al., 1980). The semantic acceptability task of Experiment 2 required that subjects recall the words of each sentence only to the point of comprehending it, not of reproducing it verbatim. Furthermore, in Experiment 2 (unlike the Mann et al., 1980 study) the groups of subjects were matched in word reading and both seemed to rely on similar phonemic codes, as evidenced by their similar decrements in performance on phonemically similar stimuli.

Thus, when provided aurally with complete phonemic representations of words, MRLs perform much like NLs on a comprehension task (Experiment 2), but markedly different on a memory task (Experiment 1). There is little doubt that the strategic demands of an explicit memory task contribute to this difference. Do strategy differences make up the whole story? This was the question addressed in Experiment 3.

Experiment 3, like Experiment 2, entailed a sentence comprehension task, but here subjects had to read the sentences to themselves, rather than listen to them. That is, subjects had to activate for themselves the relevant phonemic word codes; the codes were not provided, full blown, in an acoustic stream. In the reading condition of Experiment 3, there was no uniform adverse effect of phonemic similarity for the two groups. Data for the NLs showed trends in that direction, whereas MRLs showed no trace of a detrimental phonemic similarity effect and actually read rhymes faster than controls. While these differences were only marginally significant, the markedly different pattern of performance of the groups across the reading and listening comprehension tasks invites speculation.

One way to begin to understand the differences between Experiment 2 and 3 is to examine closely some theoretical distinctions between reading and listening. In reading, the
phonemic codes that are initially activated are probably more abstract (i.e., less complete, less fully activated) than those involved in listening, and active processing may be required to maintain them in memory if comprehension is not rapid and straightforward. In listening, however, the complete phonemic representations of words are provided in the acoustic stream; thus, listening begins with more highly activated phonemic information (on a continuum such as that proposed by Perfetti & McCutchen, 1982; see Figure 1) than does reading. Assuming similar, nonasymptotic decay functions, codes that have higher initial activation will remain above a specified threshold longer than codes with lower initial activation. (See Figure 5.) Thus, when phonemic codes are fully specified at the outset, lack of active maintenance may be less problematic.

If MRLs fail to actively maintain codes in memory as well as NLs (as suggested by Experiment 1, as well as by much of the literature on retardation), then the initial levels of activation of those codes becomes critical. During listening, phonemic word codes in memory may be fully activated for both NLs and MRLs, and even without active maintenance may remain above threshold long enough to be involved in comprehension processes. (See the decay function of Code 1 in Figure 5.) If during reading, however, the activated phonemic codes are more abstract, failure to actively maintain the code could result in its falling below threshold before comprehension processes can make use of its information. (See the decay function of Code 2 in Figure 5.)

Such a theoretical scenario could account elegantly for the results of Experiments 1, 2, and 3. In the explicit memory task of Experiment 1, NLs actively maintained phonemic representations of the word codes in memory and showed performance decrements when words
were phonemically similar. MRLs, in contrast, did not attempt to maintain word codes in memory and, therefore, were unaffected by phonemic similarities because phonemic codes dropped below threshold. In the listening task of Experiment 2, the aural presentation of the stimuli provided such complete phonemic representations that, even without active maintenance, word codes remained above threshold long enough to produce similar phonemic similarity effects for MRLs and NLs. In the reading task of Experiment 3, the more abstract phonemic codes that were activated during silent reading decayed below threshold more quickly without active maintenance. NLs maintained the codes to the extent necessary and, as a result, showed slight phonemic similarity effects. MRLs, in contrast, failed to maintain the codes in memory and showed no phonemic similarity effects. However, MRLs showed slower reading rates overall, perhaps because they had to reread words that had faded from memory.

While such a maintenance-based explanation can account for the results of these experiments, our data do not rule out the possibility that the phonemic information activated by MRLs during silent reading is not identical to that activated by NLs. There may be subtle differences in the nature of the lexical representations activated by MRLs and NLs at the subsyllabic, phonemic level, as suggested by the data of Mason (1978) and McCutchen et al. (1989). On the other hand, the analytic task employed by McCutchen et al. (1989) may, by virtue of the degree of analysis required, make sufficient memory demands that maintenance is again an issue. Experimental paradigms different from those employed here are necessary to resolve this question.

Notice, however, that the theoretical account offered in Figure 5 can also encompass, with similar mechanisms, the hypothesis that the lexical representations activated by MRLs contain phonemic codes that are somewhat impoverished compared to those of NLs. Such impoverished codes would be less fully specified and thus lower in terms of initial activation than more complete phonemic codes. Once again, assuming similar nonasymptotic decay
functions and no active maintenance, codes lower in initial activation will fall more rapidly below the threshold at which they become unavailable to comprehension processes. While largely speculative at this point, this theoretical account provides a potential mechanism for explaining the relationships among memory process, phonemic information, and reading.

**In Conclusion**

The literature on MRLs has established that these students rarely achieve a reading level beyond the fourth grade. It may be no coincidence that fourth grade is also the time when text comprehension demands begin to overshadow single-word decoding skills. The argument advanced in this study is that MRLs' lack of reading progress is closely tied to the additional comprehension demands accompanying this level of reading.

The MRLs in this study were able to read aloud words at the same grade level as their matched controls. However, when required to read and comprehend sentences, significant differences emerged. While the results need to be interpreted with caution, this study does begin to provide empirical support for the notion that MRLs' lack of reading progress beyond 4th grade competency may be related to the phonemic processing demands that accompany this level of reading. When one reads connected text, actively maintaining the phonetic features of words (while a disadvantage with the types of sentences constructed for this task) generally enhances one's memory and improves comprehension. Thus, the failure of MRLs to actively maintain phonemic information hinders their comprehension.

While a diagnosis of mental retardation does imply pervasive developmental delays, insight into the nature of these delays as they manifest themselves in specific cognitive activities like reading can help us develop more effective interventions. This study provides a starting point from which an empirically based explanation for the reading difficulties of this population may later be advanced. The critical role of speech processes and working memory in skilled
reading is widely supported in the literature on nonretarded learners. Understanding how memory codes differ between readers of all levels of intellectual skill will inform the development of effective remedial techniques for all learners.
REFERENCES


Table 1. Example word strings from Experiment 1.

**Control** (From Mann et al., 1980)
Made Punch Fruit Glass Coke

**Rhyme** (From Mann et al., 1980)
Lane Main Rain Spain Plain

**Alliterative**
Bush Pink Bloomed Buds Bit
Table 2. Example sentences from Experiment 2.

**Control** (from Mann et al., 1980)

*Acceptable version*

Tom and Bill piled books on the chair in front of the door.

*Unacceptable version*

Tom and Bill piled books on the chair in door of the front.

**Rhyme** (from Mann et al., 1980)

*Acceptable version*

Jack and Mack stacked sacks on the track in back of the shack.

*Unacceptable version*

Jack and Mack stacked sacks on the track in shack of the back.

**Alliterative**

*Acceptable version*

Bill and Bob placed bricks on the bench in back of the barn.

*Unacceptable version*

Bill and Bob placed bricks on the bench in barn of the back.
Table 3. Example sentences from Experiment 3.

**Experimental rhyme** (from Crain-Thoreson, 1991)

*Acceptable version*

Ed said that Ned just broke his sled.

*Unacceptable version*

Just his sled said that Ed broke Ned.

**Control rhyme** (from Crain-Thoreson, 1991)

*Acceptable version*

Rick said that Carl just broke his bike.

*Unacceptable version*

Joe hurt Frank's off his arm and bed jumped.

**Experimental alliterative** (from Crain-Thoreson, 1991)

*Acceptable version*

Dave told Tom to take the ten toy turtles.

*Unacceptable version*

Ten toy turtles the told Tom take Dave.

**Control alliterative** (from Crain-Thoreson, 1991)

*Acceptable version*

Jim told Bob to hold the six toy frogs.

*Unacceptable version*

To Joey asked music the shut off Mom.
Figure Captions

Figure 1. A model of a speech activation continuum, reproduced from Perfetti & McCutchen (1982).

Figure 2. Mean accuracy for the word categories for each subject group.

Figure 3. Mean accuracy for the sentence categories for each subject group.

Figure 4. Mean reading rate score for the sentence categories for each subject group.

Figure 5. Schematization of the decay function of two codes with different initial levels of activation. With similar nonasymptotic functions, codes with higher initial levels of activation will remain above a specified threshold longer than codes with lower initial activation.