The Etiology and Remediation of Phonologically Based Word Recognition and Spelling Disabilities: Are Phonological Deficits the "Hole" Story?

Richard K. Olson, Barbara Wise, Mina C. Johnson, and Jerry Ring

University of Colorado, Boulder

Chapter to be published in Foundations of Reading Acquisition. B. Blachman (Ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc..

Running Head: Phonological Deficits and Word Recognition

Mailing Address: Richard K. Olson
Department of Psychology, Box 345
University of Colorado
Boulder, Colorado 80309

Telephone: 303-492-8865 Fax: 303-492-2967 Email: rolson@clipr.colorado.edu

Acknowledgments

This work was supported in part by program project and center grants from the NICHD (HD-11681 and HD-27802), and RO1 HD-22223. The contributions of staff members of the many Colorado school districts that participate in our research, and of the twins and their families, are gratefully acknowledged. We thank Benita Blachman, Brian Byrne, Carsten Elbro, Rebecca Felton, Barbara Forman, Ingvar Lundberg, Tom Nicolson, and David Share for their comments on an earlier draft.

I. Introduction

The chapters in this volume reflect a general consensus on the central importance of phonological decoding and phoneme awareness in the etiology and remediation of reading disabilities. The year was also marked by the publication of two important theoretical papers that are consistent with this perspective (Share, 1995; Share & Stanovich, 1995). In this chapter we provide support for the importance of phonological skills from our studies on genetic factors in reading and phonological deficits and the remediation of these deficits with talking computers. However, we will question whether phonological deficits are the whole story in reading and spelling disabilities.

The question arises from two main sources. The first is the presence of individual differences in word recognition and orthographic coding skills that are independent from phonological skills. We will briefly review evidence from identical and fraternal twins on genetic and environmental contributions to this independent variance. The second reason for questioning an exclusive role for phonological deficits in reading and spelling disabilities comes from the results of several recent and ongoing remediation studies comparing good phonological training with other methods.

The main goal of the chapter is to review our own and other recent research on training in phonological awareness and decoding for children with specific reading disabilities. We will conclude that although trained improvement in phonological skills has been accompanied by strong gains in fluent word recognition, similar gains in fluent word recognition have usually also resulted from other well-designed programs, despite their lack of explicit training or comparable gains in phonological skills. This is quite puzzling, since Share (1995) and others (Jorm & Share, 1983; Share & Stanovich, 1995) have persuasively argued that phonological recoding provides a "self-teaching" mechanism that is the primary engine driving the development of printed word recognition. We will propose research on ways to improve children's use of their trained phonological skills in flexible and fluent word recognition.

Phonological skill is not the only factor influencing the development of word recognition, but there is strong evidence for its central role in the etiology of most reading disabilities. Before turning to the training studies, we will briefly review the results of behavioral-genetic analyses that reveal both common and independent genetic influences on deficits in word recognition, phonological, and orthographic skills.

II. Behavioral-Genetic Evidence on Phonological Deficits and Reading Disability

The Center for the Study of Learning Disabilities at the University of Colorado is conducting a behavioral-genetic study of reading disabilities in identical and fraternal twins. The twins are identified from school records and are invited to the laboratory for testing if either or both members of the pair have some record of a reading problem. A comparison of identical and fraternal twin-pair similarities provides estimates of the proportion of the group deficit in reading and related skills that is due to genetic factors, shared environment, and non-shared environment

(DeFries & Fulker, 1985).

Many studies have reported strong correlations between word recognition, phonological decoding (nonword reading), and phoneme awareness (e.g., deletion and manipulation of phonemes within a syllable). Similarly high correlations are found in our sample of twins with and without reading disability. Most of our twins with reading disabilities also display the commonly reported group deficit in phonological decoding and phoneme awareness when compared with younger normal twins at the same level of word recognition (Olson, Wise, Conners, Rack, & Fulker, 1989; Rack, Snowling, & Olson, 1992). This correlational evidence and structural models of prereaders' phoneme awareness and later reading skill (Wagner, Torgesen, & Rashotte, 1994) have been used to argue that phonological deficits have a significant causal role in reading disability.

Our analyses of the twin data have shown that the correlations between word recognition and phonological deficits have a significant genetic basis (Olson et al., 1989; Olson, Forsberg, & Wise, 1994). Approximately half of the group[1] deficits in word recognition, phonological decoding, and phoneme awareness are due to genetic factors, and significant genetic correlations between these deficits indicate that common genes are involved across the variables. However, there is also phenotypic variability in each measure that is not caused by common genes. Some of this variability, particularly in word recognition, is due to shared-environment influences. In addition, there are significant independent genetic influences on word recognition and phonological skills.

The clearest separation of independent genetic influences has been found between phonological decoding and orthographic coding. Our orthographic tasks assess the precision and speed of subjects' access to word-specific orthographic patterns (e.g., quickly, which is a word? rain rane; Which is an excavation? whole hole). A complete description of our orthographic measures can be found in Olson, Forsberg, Wise, & Rack (1994). Contrary to an earlier report (Olson et al., 1989), the twin data now indicate a strong genetic influence on the group deficit in orthographic coding. Part of this genetic influence is shared with phonological decoding, but an even larger part is independent (Olson, Forsberg, Gayan, & DeFries, in preparation).

Some of the independent variance in word recognition and orthographic coding is due to shared-environment differences in print exposure (Olson et al., 1994; Stanovich & West, 1989). In addition, independent genetic influence on word recognition and orthographic coding further supports our contention that phonological skills are not the only important factor in the development of word recognition. The results from remediation studies reviewed in the next section are consistent with this view.

III. Remediation of Phonological Deficits and Consequences for Word Recognition

Share's (1995) argument for the self-teaching role of phonological decoding is straightforward: Children who have relatively good phoneme awareness and phonological decoding skills should be able to decode unfamiliar words more accurately, attend to their orthographic detail, and ultimately establish more accurate and rapidly accessible orthographic codes for whole words. We have already noted that many children with reading disabilities tend to have substantial deficits in phonological decoding and phoneme awareness, suggesting a causal role in many cases of reading disability. But the evidence for a causal role is often only correlational. More convincing evidence would be provided by training studies showing that improvement in phoneme awareness and phonological decoding is followed directly by the more rapid growth of word recognition when compared to an appropriate control group with structured reading experience but no explicit instruction or comparable gains specifically in phoneme awareness and phonological decoding. We will first describe some results from our own training study that employed this basic design (Wise & Olson, 1995). Then we will review several recent studies from other laboratories showing a similar pattern of results.

The Colorado Remediation Studies

Our research on the remediation of reading disabilities began with the development of a talking-computer system to provide speech support for children's decoding difficulties (Olson, Foltz, & Wise, 1986). This system, later nick-named "ROSS" (Reading with Orthographic and Segmented Speech feedback), provided high quality synthetic speech (DECtalk) pronouncing words that children targeted with a mouse while they were reading stories on the computer screen (Olson & Wise, 1992; Wise, Olson, Anstett, Andrews, Terjak, Schneider, Kostuch, & Kriho, 1989). The subject samples included children in the 2nd-6th grades who were below the tenth percentile locally in word decoding and met the usual exclusionary criteria for specific reading disability. The studies were run in the Boulder schools over a semester time period. Trained subjects were pulled from their remedial reading or language arts class rooms for an average of 4 half-hour sessions per week. Following several sessions of pretesting and one to two weeks of training subjects to target difficult words, the subjects independently read stories on the computer for 3 of the 4 weekly sessions. The difficulty level of the stories was adjusted for each subject so that they typically would not require feedback for more than 5% of the words on a screen. The computer was programmed to present occasional multiple-choice comprehension questions and to present some of the targeted words in a review test. The experimenter monitored the subject's oral reading and encouraged targeting of unknown words during one session each week. Total training time on the computer averaged 14 hours and the time between pre and post tests averaged about 12 weeks. A randomly assigned comparison group was pre-tested and post-tested at the same times as the trained groups, but they otherwise remained with their normal reading or language arts classes.

Two main questions were addressed in our early studies with the ROSS program with children with reading disabilities. First, could reading stories on the computer with speech support improve gains in phonological decoding and word recognition more than in similar control subjects who remained in their remedial reading or language arts classes? The answer to this question was clearly yes. Trained subjects averaged about four times the gains in phonological decoding (nonword reading) and about twice the gains in word recognition on the Peabody Individual Achievement Test (PIAT) (Dunn & Markwardt, 1970).

The second main question was whether there were unique benefits from different types of orthographic and speech segmentation in the feedback provided for targeted words. We compared groups that received (1) whole-word feedback with the word highlighted as a unit and then spoken as a unit, (2) syllable feedback with syllables successively highlighted as they were spoken with a brief separation, and (3) onset-rime feedback with the onset (initial consonant or consonant cluster) and rime (vowel and following consonant(s)) of each syllable successively highlighted and spoken. An initial study suggested that gains in phonological decoding were greater with segmented feedback (Wise et al., 1989), but this effect did not replicate in a much larger sample (Olson & Wise, 1992). We have also failed to find a segmentation advantage for gains in phonological decoding or word recognition in a home-based study with 50-60 hours of reading on the computer over six months! (Olson & Wise, in preparation). We have recently found a small but significant advantage for segmentation when subjects are tested on the words they have targeted in the stories and when the colored segments appear in the tested words, possibly encouraging the children to apply their decoding strategies (Wise & Olson, in preparation).

The Olson and Wise (1992) study has been mistakenly interpreted as arguing that children with reading disabilities do not benefit from training in the analysis of sub-word orthography and speech (Lovett, Borden, DeLuca, Lacerenza, Benson, & Brackstone, 1994). However, we only concluded that there appeared to be no unique benefits for gains in phonological decoding or PIAT word recognition from segmented feedback for difficult words while reading stories. We are not surprised to learn of unique benefits from orthographic and speech segmentation while practicing a specific list of trained words or of improved generalization to similar words (Lovett, Barron, Forbes, Cuksts, & Steinbach, 1994). In fact, the segmental analysis of words and the training of phonological awareness has been a central focus of our subsequent remedial studies (Wise & Olson, 1995).

Olson and Wise (1992) noted that subjects who had the lowest initial levels of phoneme awareness (adjusted for initial level of word recognition) tended to gain significantly less from reading with speech feedback, and there was a marginally significant interaction suggesting that these subjects' gains were particularly low when their feedback was segmented into the smallest units (onset-rime). That result, along with our own and others' evidence for disabled readers' group deficit in phoneme awareness and phonological decoding, led us to focus our next group of studies on improving phonological skills prior to and concurrently with reading stories on ROSS (Wise & Olson, 1995). Our hope was that such improvement would advance the self-teaching mechanism hypothesized by Share (1995) and accelerate the development of subjects' word recognition.

The method we chose to improve Phoneme Awareness (PA) and phonological decoding was our adaptation of the Auditory Discrimination in Depth program (Lindamood and Lindamood, 1975). Briefly, the PA program begins by helping children to discover the distinctive articulatory movements associated with different speech sounds. For instance, the first session included motivation and introduction of the program's first three "quiet and noisy pairs": "lip poppers /p,b/, tip tappers /t,d/, and back scrapers /k,g/." In small groups consisting of the trainer and three students, children used mirrors to help them discover what their mouths were doing for different phonemes and they learned to associate appropriate mouth pictures, articulatory labels, and letters with those phonemes. They practiced these associations on the computer using programs under development at the Lindamood-Bell Company in California. They then spent 20 minutes in small groups ordering a sequence of mouth pictures to indicate the sequence of phonemes in a syllable. After that, they next used programs developed at the University of Colorado to develop phonological awareness and phonological decoding with letter symbols. In one program, subjects practiced sequencing and manipulating the letter-symbols to correspond to changes in syllables spoken by the computer. In the second program, children explored English orthographic patterns and print-to-sound relations in a spelling exploration game where the computer pronounced correct and incorrect attempts, and children could modify their attempts (Wise & Olson, 1992). In the third program, children selected printed nonwords to match nonwords spoken by the computer.

When students achieved 80% success over 2 days manipulating sounds and symbols in CVC words, they began spending half their computer time reading on ROSS. The ROSS speech support for targeted words was segmented at the onset-rime level for regular one-syllable words and at the syllable level for multi-syllable words. (Exception words such as "said" were not segmented.) The subjects had to push the mouse button first to highlight each segment in different colors. They attempted to sound out the segments with help from the articulatory and phonics (letter-sound) knowledge they had developed from earlier training. Then they pressed another button to have the computer successively highlight and speak each segment. A third button could optionally be pressed last to get whole-word support if needed. The total training time of 25 hours consisted of 7 hours in small-group sessions learning the articulatory and phonics concepts, 10 hours in practice on related computer games, and 8 hours reading stories on ROSS.

A trained comparison group had an equal amount of small-group interaction with the trainer (7 hours), but their initial small-group sessions were focused on methods for developing the Comprehension Strategies (CS) of predicting, generating questions, clarifying, and summarizing. This approach has shown positive effects for improving reading comprehension in groups of poor comprehenders (Palinscar & Brown, 1984). Children in the CS group began reading stories from the beginning of training while the PA group was occupied with practice in articulatory awareness and the manipulation of phoneme and letter sequences.

Our main goals for the CS comparison group were first, to equate the level of group interaction, enthusiasm, and trainer involvement, and second, to provide greater print exposure with speech and segmentation support for decoding assistance in stories. We wanted to contrast the effects of the phonological awareness and decoding training in the PA group while controlling also for teacher and school effects by having both methods taught by all teachers in all schools. The CS group spent 7 hours in the initial small-group sessions reading and discussing stories, and

18 hours reading stories on ROSS. Decoding and speech support for targeted words was segmented in the same way as in the PA group, and subjects were told they should sound out each segment before asking for speech support to help them read targeted words. The CS subjects varied widely in the degree to which they seemed to use the segments to aid decoding.

The results from the two training conditions showed significant gains in phoneme awareness and phonological decoding in both groups. The CS group made gains in these skills that were stronger than in previous ROSS studies, suggesting that the engaging strategies and the daily interaction with a teacher benefitted both groups. The most important result for our theoretical discussion was that the PA group made much greater gains than the CS group in phoneme awareness (three times greater) and phonological decoding (nearly two times greater).

Based on Share's (1995) self-teaching hypothesis, we had expected that the PA group's substantially greater gains in phoneme awareness and phonological decoding would also yield gains in word recognition that were substantially greater than those in the CS group. Both PA and CS groups made gains in word recognition that were larger than in previous ROSS studies. However, the groups were not significantly different on standardized tests of untrained words, despite their strong differences in phonological awareness and decoding. There was a nonsignificant trend towards a PA group advantage on the PIAT word recognition test (gains of 1.1 vs. .9 grade levels). The PA group was significantly worse in an experimental test of word recognition that provided a time-limited two seconds of exposure for each word in the list. There were no significant group differences in gains on reading comprehension or spelling.

The PA group did show significant advantages in word recognition performance on all tests of subsets of previously targeted words from the ROSS stories. These tests were given at the end of each monitored session, after each month using words targeted during the month, and at the end of training for a subset of words targeted over the semester. However, the words in these tests were orthographically segmented as they had been when originally targeted in the ROSS stories. It is not clear if the PA group's better performance was due to better initial encoding of the word when it was first targeted in the story, better decoding on the test, or both. In any case, the explicit task demands on the PA group during the initial encoding and testing of targeted words were unlike our pre- and post-test measures of word recognition. In these measures, unsegmented words were presented one at a time on the computer screen. Apparently, the students in the PA group tended not to use their improved phonological decoding skills in this more natural presentation of tested words.

Wise and Olson (1995) speculated that the PA subjects may not have had enough print exposure to integrate their superior phonological skills in reading and build the larger reading vocabularies needed to surpass the CS group's gains on our measures of word recognition. We also suggested that the PA group's significantly inferior performance on the time-limited word recognition test was caused by their slower and more analytic decoding approach. Following more print exposure and opportunity to automatize their phonological decoding, we suggested that the PA group would equal or surpass the CS group's gains in time-limited word recognition. We hoped that these predictions, based on the self-teaching hypothesis, would be confirmed by our planned longitudinal follow-up testing.

A Brief Comment on Results for Reading Comprehension

Before turning to our unpublished follow-up results, we will briefly comment on the null results for group differences in reading comprehension. Children enjoyed the Comprehension Strategy (CS) training, with its high level of social interaction and discussion of stories. While we achieved our goal of balancing the general level of group interaction and student motivation across conditions, we had also hoped to find a significant advantage for students' reading comprehension on the PIAT and the Gates MacGinitie Reading Comprehension of Passages (MacGinitie and MacGinitie, 1989). Surprisingly, nonsignificant trends in comprehension gains actually favored the PA group. The only measure indicating a significant comprehension advantage for the CS group was the percent correct on ROSS comprehension questions (90% PA vs 95% CS) that were answered in sessions when subjects' oral reading was monitored by a trainer who promoted the CS subjects' use of comprehension strategies. There was no significant difference in sessions where subjects read silently without the trainer (both 86% correct). Apparently the CS subjects did not spontaneously use their trained comprehension strategies when reading on their own. This suggests a transfer problem similar to that observed on our post-test measures of word recognition for the PA subjects, who may not have consistently applied their improved phonological decoding skills.

Johnson (1995) asked a subset of the Olson and Wise (1995) PA and CS subjects to give oral summaries of stories read silently as pre- and post-tests. The CS subjects showed significantly better main idea representation and less intrusion of irrelevant details. (The CS subjects were explicitly trained to summarize main ideas.)

Results from a Long-Term Follow-up to Wise and Olson (1995)

No group differences in rate of gain.

About nine months (including summer vacation) after the end of training in the Wise and Olson (1995) study, subjects were retested on phoneme deletion, nonword reading, PIAT word recognition, and Time-Limited Word Recognition. Ring (1995) analyzed the data for the pre-test, mid-test, post-test, and final nine-month follow-up test. Figure 1 a, b, c, and d show the results for 47 subjects (PA, N = 27; CS, N = 20) from three of the four schools in the original study. The fourth school was excluded because they received extra training in other experimental methods during the following year. In the other three schools, subjects maintained their advantage in phonological decoding (Figure 1 a) and slightly less so in phoneme awareness (Figure 1 b). However, the PA subjects still did not show significantly greater gains in either PIAT word recognition (Figure 1 c) or in time-limited word recognition (Figure 1 d), although the PA group was no longer significantly worse in time-limited word recognition.

Insert Figure 1 a - d about here

Drop-off in rate of gain following training.

It is clear in Figures 1 c and d that the rate of gain in word recognition over the nine months following training was substantially less than the rate during training. Ring (1995) found a significant quadratic function that reflected the decline in growth rate following training. It appears that a continuation of structured reading practice similar to the ROSS program would be necessary to maintain the initial high rate of improvement observed during training. Subjects' print exposure during the 9 months following ROSS training may have been too limited and/or it lacked the corrective speech support subjects needed to decode difficult words and add them to their word-recognition vocabulary.

We have some evidence from an earlier unpublished study showing that the speech support provided while reading on ROSS and our emphasis on accurate word recognition is essential for producing the strong gains observed during training. Children were randomly assigned to a group that read on ROSS with speech feedback for targeted words and a group that was encouraged to target difficult words and try to read them with help from the story context, but they received no speech support for unknown words. This type of independent reading was similar to what many of the children were doing with books in their "whole-language" oriented remedial reading or language arts class rooms. Unfortunately, the study could not be completed because we had considerable problems motivating children in the no-feedback condition to continue with the training. The limited post-test data that was collected revealed minimal gains from the no-feedback condition.

Why was there no support for the self-teaching hypothesis?

Why have we not found support for the self-teaching hypothesis (Share, 1995) in the PA group with its significantly greater long-term gains in phoneme awareness and phonological decoding? There are several possible answers. Perhaps we had not trained subjects' phonological skills to a high enough level of accuracy, although their accuracy in nonword reading and phoneme awareness approached that of normal readers at their age level. Perhaps our encouraging the PA subjects' use of phonological decoding during their 8 hours of reading on ROSS was not enough to promote transfer of their new phonological skills to reading. It has been noted that learning disabled children often do not use the strategies they have been successfully trained on (Borkowski, Estrada, Milstead, & Hale, 1989; Kavale, 1980). Later we will continue this line of argument and discuss some other possible reasons for transfer failure. First we will review some recent studies by other researchers that have also failed to find significantly greater gains in word recognition despite strong gains in phonological coding and phoneme awareness following successful phonological training, compared to well-trained reading-only control groups.

Other Studies with Similar Null Results for Word Recognition

The phonological training studies reported in this volume all seem to tell the same story across widely varying subject populations. Byrne's (this volume) training of phonological awareness in unselected preschoolers led to greater gains three years later in nonword reading and surprisingly in reading comprehension, but not generally in word recognition. Byrne did find that there was a significant advantage for the trained group on 5 of the least frequent words in their regular word list. Torgesen's (this volume) articulatory-based phonological training of 52 unselected children from the second half of kindergarten through the first grade led to better nonword reading but not significantly better word recognition, reading comprehension, or spelling at the end of first grade and in the second grade, compared to an implicit phonics group. Foorman's (this volume) comparison of remedial teaching methods (potentially confounded by school and teacher) for 108 second- and third-grade children with reading disabilities included synthetic/analytic phonics, synthetic phonics, and the Edmark sight-word approaches. After controlling for group differences in memory and demographics, growth-curve analyses indicated that the synthetic phonics approach was most effective in training phonological processing, but the advantage in phonological processing did not lead to greater gains in word recognition when compared to the Edmark program. Finally, Nicolson (this volume) reviewed his remediation studies that showed greater gains in phonological skills (nonword reading) but no corresponding advantages in word recognition or passage reading.

Nicholson (personal communication, 1996) hypothesized that his phonologically trained subjects' advantage in nonword reading will eventually lead to better word recognition, but the follow-up data needed to support this hypothesis are not yet available. Nicholson also suggested that since word recognition tests typically contain many irregular words, they may not be a fair measure of subjects' gains in word recognition. This may be a valid criticism if the measures contain a disproportionate number of exception words compared to their distribution in the childrens' written language. We are now assessing this possibility in our own measures.

There were non-significant trends in an interesting study by Lovett et al. (1994) suggesting that a letter-sound phonics approach may result in slightly better reading of simple regular words and poorer reading of exception words in children with severe reading disabilities. We are also interested in this study because both the authors of the Lovett et al. study and Share and Stanovich (1995) describe the results as providing clear evidence for the unique benefits of phonological training in the development of word recognition. However, a closer look at the Lovett et al. study actually reveals results for treatment differences in word recognition that are similar to those of Wise and Olson (1995).

Lovett et al. (1994) trained three groups of 7-13 year old children with severe reading disabilities from a clinic population, in a clinic setting, for a total of 35 one-hour sessions. Two of the treatment conditions were designed to promote the development of word decoding skills but in different ways. A Phonological Analysis and Blending/Direct Instruction (PHAB/DI) program (similar to Engleman and Bruner, 1988) focused on letter sound correspondences and the segmentation and blending of phonemes in a corpus of regular words. A second treatment condition, the Word Identification Strategy Training (WIST) program, was based on part of a "metacognitive phonics" program developed by Gaskins and Elliot (1991). This program uses a "whole-word" approach to train 5 key-words each day (120 total) that contain a range of high-frequency spelling patterns. (These were the same words used in the PHAB/DI program.) Children were taught to use their knowledge of the spelling patterns in the key words to decode by analogy the parts of new words that contained similar spelling patterns. Other strategies included trying different vowel pronunciations and the "peeling off" of prefixes and suffixes to first decode the root word. Both methods promoted sub-word decoding skills, but the PHAB/DI program was more oriented toward the segmenting and blending of phonemes and their corresponding letters, while the WIST program emphasized larger spelling patterns and decoding by analogy.

Lovett et al. (1994) trained a third "control" group in a variety of organizational and study skills to control for simple "Hawthorn effects". However, it appeared that this training may have also displaced reading instruction that subjects would otherwise have had in the clinic. Thus, it is not surprising to learn that both of the trained groups gained significantly more than the control group on most reading measures, except reading comprehension. A control condition of reading with feedback but without special decoding instruction would have been needed to assess differences between decoding instruction and accurate reading practice. Nevertheless, a comparison of results from the two decoding programs in Lovett et al. can be used to make the same point.

When compared to the WIST program, the PHAB/DI program resulted in significantly greater improvement on phonological awareness and phonological decoding by the end of training. Unfortunately, as in our own study and in the other studies reviewed above, the PHAB/DI group's better phonological performance did not generally lead to better word recognition compared to the WIST group. The Lovett et al. (1994) study was remarkably well endowed with a broad array of word recognition measures. The pre- and post-test measures included the 120 trained key words, a list of 371 words similar to the key words, 105 multisyllabic words with embedded key-word spelling patterns, 298 regular words, 298 exception words, and 10 difficult low-frequency words for a grand total of 1,202 test words! Standardized measures of word recognition included the Wide Range Achievement Test-Revised (WRAT-R)(Jastak & Wilkinson, 1984), and the Woodcock Reading Mastery Test-Revised (WRMT-R) (Woodcock, 1987). On all but two of the measures, there were no significant training group differences. The WIST group did significantly better on the 10 difficult words, and the PHAB/DI group did significantly better on the WRAT-R test.

Lovett et al. (1994) noted that the WRAT-R test allowed 10 seconds per word and this may have given the PHAB/DI subjects the time needed to use their phonological decoding skills on the mostly regular words in this test. A group difference in response to word regularity was also suggested by trends in performance on the 298 regular and 298 exception word lists. The PHAB/DI group was slightly but nonsignificantly better on regular words while the WIST group was nonsignificantly better on the exception words. Overall, the groups' average performance in word recognition was remarkably similar, in spite of the significant group differences in phonological skills. As in the Wise and Olson (1995) study and other studies reviewed earlier, there was a dissociation between growth in phonological skills and growth in word recognition that was inconsistent with Share's (1995) self-teaching hypothesis.

A Closer Look at Two Studies Showing Greater Gains in Word Recognition

Our review of the recent literature included two other studies that are frequently cited for showing a unique contribution from phonological training to word recognition for children with reading disabilities. The first study by Brown and Felton (1990) identified children in kindergarten at risk for reading disability because of performance below the 16th percentile on three research measures of phonological processing. At the beginning of first grade, the children were randomly assigned to a "structured phonics" group (Lippincott Basic Reading Program, 1991) or a "context" group (Houghton Mifflin, 1986) for their reading instruction in the first and second grades. At the end of first grade, the structured phonics group was significantly better in nonword reading and spelling of regular words, but not in word recognition. At the end of second grade, there was still no significant difference in the standardized word recognition test given at the end of the first grade, and the previous differences in nonword reading and spelling were no longer significant. However, new measures included at the end of second grade revealed a significant advantage for the "structured phonics" group in monosyllable nonwords and polysyllable real words, but not monosyllable real words.

Felton (personal communication, 1996) collected follow up data on a subset of the above subjects at the end of eighth grade. Some additional intervention of the "code" type had been given during the third grade to four of the 17 "code" and 6 of the 13 "context" subjects who were having the most trouble reading. This additional code instruction may have worked against any reading difference in the 8th grade, and there were no significant differences on the reading test. For example, grade equivalents on a composite based on the Woodcock Johnson word identification and passage comprehension tests were 5.6 GE (N = 17) for the "code" group, and 5.3 GE (N = 13) for the context group. However, the general efficacy of <u>both</u> groups' intense small-group instruction was shown by the significantly lower performance in the 8th grade for a passive control group (3.7 GE, N = 17)

We accept the authors' contention that <u>trends</u> in the data nearly always favored the "structured phonics" group and that this is probably the superior method for these children. But we have two methodological concerns. First, the training methods were confounded by teacher, making it difficult to dissociate method effects from teacher effects. Second, the "context" group may have made many uncorrected decoding errors during their reading practice because word identity is often incorrectly predicted from context (Gough, Alford, & Holly-Wilcox, 1981). These uncorrected errors would reinforce inappropriate associations between print and speech. We will return to this issue at the end of the chapter and argue that a more appropriate comparison group would include reading with corrective feedback for decoding errors.

A second study by Hatcher, Hulme, and Ellis (1994) was the only one we could find that appeared to show a clear advantage in several measures of word recognition from reading instruction that emphasized phonological analysis. A total of 124 seven-year-old children with reading disabilities were assigned to supplemental training groups (40, 30 minute sessions over 20 weeks) that included either phonological-analysis-plus-reading instruction, reading instruction alone, or phonological analysis alone. A fourth control group received no supplemental instruction. The statistical analysis was limited because it only compared the trained groups against the untrained control group and not against each other. However, the phonological-analysis-plus-reading group was the only one that showed significantly greater gains compared to controls across all three measures of word recognition. These reading advantages were obtained in spite of the fact that the phonological-analysis-plus-reading group and the reading-only group were not significantly different from each other in their gains on four measures of phonological processing (sound deletion, nonword segmentation, sound blending, and sound categorization). Another concern is that the reading-only group without phonological training generally showed little or no reading advantage over the untrained control group. Other studies typically do show significant advantages for all reading instruction groups versus untrained groups (c.f., Lovett et al., 1994; Olson & Wise, 1992). Thus, we are concerned that there may have been some problem with the reading comparison group beyond the absence of phonological training. The lack of significant group differences in phoneme awareness after phonological training raises questions about the contribution of this training to better word recognition in the phonological-analysis-plus-reading group.

Studies Showing Generalization to Word Recognition for Unselected Samples

There have been a few recent studies with unselected samples of preschoolers, kindergarten, or first-grade beginning readers that have suggested generalization from better phonological skills to better word recognition. Lundberg, Frost, & Peterson, (1988) reported significantly better word recognition in Danish children who had received substantial training in phonological awareness, without print, as 6 year old prereaders. The significant advantage in word recognition emerged in the second grade but was not present in the first grade. Of particular relevance to children at risk for reading disability, the benefits were significantly stronger in those children whose initial phonological skills were unusually low (Lundberg, 1994). This classic study is unique in showing the effects of phonological training without the inclusion of print. It has a potential confound of different teachers and different Danish islands for the treated and control groups, although the authors carefully tried to equate the samples on relevant demographic variables.

Uhry and Shepherd (1993) reported a study with unselected first graders that showed strong gains in word recognition and phonological awareness. They employed exercises in manipulating phonemes and letters in spelling analysis and production as a key element in their experimental training program. Children (N = 11) in this group also used several computer games to foster the development of sub-word analysis in reading. A trained comparison group (N = 11) spent more of their time working with print at a whole-word level of analysis, although they were encouraged to use decoding skills that were taught to both groups in the regular classroom. In spite of the small samples and some common elements of phonological decoding training for both groups in the regular classroom, the spelling group showed significantly greater gains in phonological awareness, phonological decoding, and word recognition. The small samples of children in the Uhry and Shepherd study were from college educated families, most attended 3 years of preschool prior to kindergarten, and their mean IQ score was 122.5. It would be very interesting to see if the unique benefits of their spelling and sound manipulation program would also generalize to a reading disabled sample and to beginning readers at risk for reading disability.

Ball and Blachman (1991) compared three groups of kindergarten children from inner-city schools. The two trained groups of interest here were given 28 20-minute sessions of either letter-name-sound training and general language activities (Language Activities group), or letter-name-sound training plus exercises in manipulating phonemes within words represented by tiles and ultimately individual letters (only one per word) within words (Phoneme Awareness group). By the end of training, the Phoneme Awareness group improved significantly more on phoneme segmentation and on the mean number of words read correctly on the Woodcock Word Identification Subtest (3 words versus 1 word). A second kindergarten study compared phonological training with an untrained control group. The trained group was significantly better on post-tests of phonological skills and regular word reading, but not on the Woodcock standardized measure of word recognition (Blachman, Ball, Black, & Tangel, 1994). During the first grade, the trained group participated in a reading program that emphasized the alphabetic code, while the control children participated in a basal reading program. At the end of first grade, and in a follow-up at the end of second grade, the trained group significantly outperformed the control group on measures of regular word reading. Differences between the two groups on the Woodcock standardized measure of word recognition also favored the trained children. These differences were marginally significant ($p = .056$) at the end of grade one and significant at the end of grade two (Blachman, Ball, Black, & Tangel, in preparation).

The above studies with unselected samples suggest that related advantages in phonological processing and word recognition can be found when comparing different methods of beginning reading instruction. A similar consensus has emerged from reviews of earlier studies of different training programs for beginning readers (Adams, 1990; Chall, 1983), although Adams cautions that other confounding variables such as teacher enthusiasm for a new and favored method or different time on task may have influenced the results of some studies. It appears that the studies reviewed in this section were carefully controlled, but further research is needed to show that their methods can provide unique benefits for phonological skills and word recognition among children with reading disabilities.

IV. Summary and The Way Ahead

The studies of children with reading disabilities reviewed in the last section should not be used to argue the null hypothesis against any additional gains in word recognition from training in phonological skills. Sometimes there were nonsignificant trends favoring phonological training. Sometimes a significant advantage was found when task demands and/or the lack of time constraints seemed to favor the use of slow and deliberate phonological decoding for regular words, as in the Lovett et al. (1994) study and our most recent training study. In our second training study with the PA and CS conditions (Wise & Olson, in preparation), a significant post-test advantage was found for PA subjects' gains in the WRAT-R test of word recognition, but not for gains on the PIAT test, our experimental measure of time-limited word recognition, or miscues on oral reading of paragraphs. Earlier we noted similar results from Lovett et al. (1994). They reported that the liberally timed and largely regular-word WRAT-R test was the only one of their numerous word recognition measures that showed a significant advantage for the phonologically trained group. The question we turn to now is how to increase the effects of phonological training on the development of word decoding that is both fluent and flexible enough to deal with the irregularities of English orthography in normal reading.

Share (1995) may be correct in his hypothesis that normally developing readers do use their superior phonological skills as a primary self-teaching mechanism to develop more accurate, fluent, and flexible word recognition. However, this mechanism may continue to be compromised in children with initially very weak phonological skills, even after training has produced better phonological awareness and decoding at the conscious level of processing required for greater accuracy in our measures (Bowey, 1995). Subjects' performance on our measures may provide only a rough index of the underlying phonological processes that are employed very rapidly and automatically (not requiring much attention) in normal readers' fluent word recognition (Perfetti & Bell, 1991; Van Orden, 1987). Training children with reading disability in conscious and deliberate phonological awareness and decoding may not automatically result in phonological processing that is fast and automatic enough to aid in the development and execution of fluent word recognition. If this hypothesized dissociation between levels of phonological processing is correct, then we may need to focus our efforts on training phonological processes beyond accuracy to a higher level of speed, automaticity, and flexibility.

One approach we are currently exploring is longer training time following the same PA training procedures and reading on ROSS used in Wise and Olson (1995). Recall that the PA group in Wise and Olson averaged only 8 hours reading stories on ROSS. We are currently training a small sample of children with reading disabilities over a full year and hope to reach a total of 40 hours training time, including at least 15 hours on ROSS, by the end of the spring semester. This longer training time should provide some additional boost in performance on our phonological measures. More importantly, it will provide much more time for the PA subjects to automatize and integrate their phonological processes with their reading of stories with speech support in the ROSS program.

For next year's study, phonological training will be followed by new computer exercises that encourage greater flexibility in trying different vowel pronunciations and stress patterns with real word reading. Additional speed training will promote the rapid identification of printed words, particularly high frequency exception words, and nonwords spoken by the computer.

We are optimistic that the above procedures will result in more fluent and flexible word recognition for most children with reading disabilities when compared to similar children who receive CS training. We also hope that the self-teaching function of improved accuracy, speed, and automaticity of the PA group's phonological skills will result in more rapid growth in word recognition long after training. However, as our title implies, we do not think that phonological deficits are the whole story. The development of printed word recognition is also dependent on other factors such as general print exposure (Stanovich & West, 1989; Olson et al., 1994), accurate reading with speech support for decoding errors by person or computer (Olson & Wise, 1992; Wise & Olson, 1995), genetically-based individual differences in orthographic skills that are independent from phonological skills (Olson et al., 1994; in preparation), lexical knowledge (Morton, 1989; Plaut, McClelland, Seidenberg, & Patterson, 1996), the use of context (Tunmer and Chapman, 1995), and speed of lexical access (Bowers & Wolf, 1993; Wolf, this volume). The effects of these factors on the growth of printed word recognition may often be largely independent from what we do for deficits in phonological processing.

We conclude with a critically important point for theory and practice in remediation: Our CS condition with its comprehension strategy training and reading on the computer was not like some "whole-language" approaches that encourage guessing based on context (c.f., Houghton & Mifflin, 1986). Gough et al. (1981) and others have pointed out that guessing from context often yields errors in word decoding that can weaken the appropriate connections between orthography and speech (Jorm & Share, 1983). Instead, our CS subjects were trained to attempt all words and to use speech support when they were unsure of a word. We believe that this support and attention to correct decoding is necessary for the rapid development of accurate and fluent word recognition. Also, it is consistent with the essential role of corrective feedback in modern connectionist models of reading development (Adams, 1990; Plaut et al., 1996; Seidenberg & McClelland, 1989).

Children with reading disabilities usually require more trials with feedback than normal readers to learn the correct decoding for unknown words and more repeated exposures to develop fluent recognition (Reitsma, 1983; Ehri & Saltmarsh, 1995). Supportive parents and teachers may be the ideal source for this additional corrective feedback, but their availability is often limited. We have shown that our ROSS computer programs can provide the additional feedback and print exposure that is needed to substantially improve fluent and flexible word recognition, especially when used in a well structured and motivated setting. A remaining challenge is to demonstrate whether and how intensive phonological training can significantly improve on these gains.

References

Adams, M.J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Ball, E.W., & Blachman, B.A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? Reading Research Quarterly, 26, 49-66.

Blachman, B.A., Ball, E.W., Black, R.S., & Tangel, D.M. (1994). Kindergarten teachers develop phoneme awareness in low-income, inner-city classrooms: Does it make a difference? Reading and Writing: An Interdisciplinary Journal, 6, 1-18.

Borkowski, J.G., Estrada, M.T., Milstead, M., & Hale, C.A. (1989). General problem solving skills relations between metacognition and strategic processing. Learning Disability Quarterly, 12, 57-70.

Bowers, P.G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. Reading and Writing, 5, 69-85.

Bowey, J.A. (1995). On the contribution of phonological sensitivity to phonological recoding. Issues in Education, 1, 65-70.

Brown, I.S., & Felton, R.C. (1990). Effects of instruction on beginning reading skills in children at risk for reading disability. Reading and Writing: An Interdisciplinary Journal, 2, 223-241.

Cardon, L.R., Smith, S., Fulker, D., Kimberling, W., Pennington, B. & DeFries, J. (1994). Quantitative trait locus for reading disability on chromosome 6. Science, 266, 276-279.

Chall, J.S. (1983). Learning to read: The great debate (2nd ed.). New York: McGraw-Hill.

DeFries, J.C., & Fulker, D.W. (1985). Multiple regression analysis of twin data. Behavior Genetics, 15, 467-473.

Dunn, L.M., & Markwardt, F.C. (1970). Examiner's manual: Peabody Individual Achievement Test, Circle Pines, MN: American Guidance Service.

Ehri, L.C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. Reading and Writing: An Interdisciplinary Journal, 7, 295-326.

Engelmann, S., & Bruner, E.C. (1988). Reading Mastery I/II Fast Cycle: Teachers guide. Chicago: Science Research Associates.

Gaskins, I.W., & Elliot, T.T. (1991). Implementing cognitive strategy training across the school: The Benchmark manual for teachers. Cambridge, MA:Brookline Books.

Gough, P.B., Alford, J.A., Jr., & Holley-Wilcox, P. (1981). Words and contexts. In O.J.L. Tzeng & H. Singer (Eds.), Perception of print (pp. 85-102). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hatcher, P.J., Hulme, C., & Ellis, A.W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. Child Development, 65, 41-57.

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

Johnson, M.C. (1995). Effects of training in phonological awareness and reciprocal teaching on the comprehension of disabled readers. Unpublished Masters Thesis. University of Colorado, Boulder.

Jorm, A.F., & Share, D.L. (1983). Phonological recoding and reading acquisition. Applied Psycholinguistics, 4, 103-147.

Kavale, K.A. (1980). The reasoning abilities of normal and learning disabled readers on measures of reading comprehension. Learning Disability Quarterly, 3, 34-45.

Lindamood, C. & Lindamood, P. (1975). Auditory Discrimination in Depth. Science Research associates Division, MacMillan/McGraw Hill: Columbus: OH.

Lippincott Basic Reading 1981. Riverside, NJ: Macmillan Publishing Company.

Lovett, M.W., Barron, R.W., Forbes, J.E., Cuksts, B., & Steinbach, K.A. (1994). Computer speech-based training of literacy skills in neurologically-impaired children: A controlled evaluation. Brain and Language, 47, 117-154.

Lovett, M.W., Borden, S.L., DeLuca, T., Lacerenza, L., Benson, N.J., & Brackstone, D.

(1994). Treating the core deficits of developmental dyslexia: Evidence of transfer of learning after phonologically- and strategy-based reading training programs. Developmental Psychology, 30, 805-822.

Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in pre-school children. Reading Research Quarterly, 23, 263-284.

Lundberg, I. (1994). Reading difficulties can be predicted and prevented: A scandinavian perspective on phonological awareness and reading. In C. Hulme & M. Snowling (Eds.), Reading development and dyslexia (pp 180-199). London: Whurr Publishers.

MacGinitie, W. and MacGinitie, R. (1989). The Gates-Macginitie Reading Tests (3rd ed.). Riverside Publishing: Chicago, IL.

Morton, J. (1989). An information processing account of reading acquisition. In A.M. Galaburda (Ed.). From reading to neurons. Cambridge, Mass: MIT Press.

Olson, R. K., Foltz, G., & Wise, B. Reading instruction and remediation with the aid of computer speech. Behavior Research Methods, Instruments, and Computers, 1986, 18, 93-99.

Olson, R.K., Forsberg, H., & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V.W. Berninger (Ed.), The varieties of orthographic knowledge I: Theoretical and developmental issues (pp. 27-71). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G.R. Lyon (Ed.) Frames of reference for the assessment of learning disabilities: New views on measurement issues (pp. 243-277). Baltimore: Paul H. Brookes Publishing Co.

Olson, R.K. & Wise, B.W. (1992). Reading on the computer with orthographic and speech feedback: An overview of the Colorado Remedial Reading Project. Reading and Writing: An Interdisciplinary Journal, 4, 107-144.

Olson, R.K., Wise, B., Conners, F., Rack, J., and Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. Journal of Learning Disabilities, 22, 6, 339-348.

Palincsar, A.S. & Brown, A.L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activity. Cognition and Instruction, 2, 117-175.

Perfetti, C.A., & Bell, L. (1991). Phonemic activation during the first 40 ms. of word identification: Evidence from backward masking and masked priming. Journal of Memory and Language, 30, 473-485.

Plaut, D.C., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in Quasi-Regular Domains. Psychological Review, 103, 56-115.

Rack, J.P., Snowling, M.J., & Olson, R.K. (1992). The nonword reading deficit in developmental dyslexia: a review. Reading Research Quarterly, 27(1), 28-53.

Reitsma, P. (1983). Printed word learning in beginning readers. Journal of Experimental Child Psychology, 75, 321-339.

Ring, J. (1995). A comparison of computer-based remedial reading programs: The effects of automaticity, phonological awareness, and comprehension strategy training on word recognition and phonological decoding skills. Unpublished Masters Thesis. University of Colorado, Boulder.

Seidenberg, M.S., & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. Psychological Review, 96, 523-568.

Share, D.L. (1995). Phonological recoding and self-teaching: sina qua non of reading acquisition. Cognition, 55, 151-218.

Share, D.L., & Stanovich, K.E. (1995). Cognitive processes in early reading development: Accommodating individual differences into a model of acquisition. Issues in Education, 1, 1-57.

Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. Reading Research Quarterly, 24, 402-433.

Tunmer, W.E., & Chapman, J.W. (1995). Context use in early reading development: Premature exclusion of a source of individual differences? Issues in Education, 1, 97-100.

Uhry, J.K., & Shepherd, M.J. (1983). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. Reading Research Quarterly, 28, 219-233.

Van Orden, G.C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. Memory and Cognition, 15, 181-198.

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. Developmental Psychology, 30, 73-87.

Wise, B.W. and Olson, R.K. 1992. Spelling exploration with a talking computer improves phonological coding. Reading and Writing, **4**:145-156.

Wise, B.W., & Olson, R.K. (1995). Computer-based phonological awareness and reading instruction. Annals of Dyslexia, 45, 99-122.

Wise, B., Olson, R., Anstett, M., Andrews, L., Terjak, M., Schneider, V., & Kostuch, J. (1989). Implementing a long-term computerized remedial reading program with synthetic speech feedback: Hardware, software, and real-world issues. Behavior Research Methods, Instruments, and Computers, 21, 173-180.

Woodcock, R.W. (1987). Woodcock Reading Mastery Tests-Revised. Circle Pines, MN: American Guidance Service.

Footnote

The emphasis on the group deficit is to indicate that the present behavioral-genetic analyses do not provide information on the proportion of genetic and environmental influence for any individual within the group. Rather, we are only estimating the average proportions for genetic and environmental influences across the group. More specific statements about genetic influence at the individual level may ultimately be possible, for example, when a gene linked to reading disability on chromosome 6 is finally identified (Cardon et al., 1994).

Figure Captions

Figure 1a: Gains in phonological decoding, percent correct.

Figure 1b: Gains in phoneme deletion, percent correct.

Figure 1c: Gains in PIAT word recognition, items correct.
Figure 1d: Gains in time-limited word recognition, items correct.