Published in final edited form as:
Sci Stud Read. 2010 January ; 12(1): 90–105. doi:10.1080/10888430701773876.

Phonological Spelling and Reading Deficits in Children with Spelling Disabilities

Angela Friend and **Richard K. Olson**
University of Colorado, Boulder

Abstract

Spelling errors in the Wide Range Achievement Test were analyzed for 77 pairs of children, each of which included one older child with spelling disability (SD) and one spelling-level-matched younger child with normal spelling ability from the Colorado Learning Disabilities Research Center database. Spelling error analysis consisted of a percent graphotactic-accuracy (GA) score based on syllable position and existence in English, and a phonological accuracy score (PA). The SD group scored significantly worse in the PA measure, and non-significantly better than controls on the GA measure. The group by measure interaction was significant. Spelling matched pairs had very similar scores for word recognition and orthographic coding, but the SD group exhibited significant deficits in reading measures of phonological decoding and in language measures of phonological awareness.

Many studies have found robust correlations between disabilities in word reading, phonological decoding of printed nonwords, and language measures of phonological awareness. Moreover, the potential causal role of phonological deficits in reading disabilities has been supported by "reading-level-match" studies, wherein older children with reading disabilities are compared to younger normally developing readers at the same absolute level of word recognition. Most such studies of English readers have reported significantly lower phonological decoding and/or phonological awareness for children with reading disabilities (for review, see Rack, Snowling, & Olson, 1992; Vellutino et al., 2004).

Poor phonological decoding and poor phonological awareness are also correlated with poor spelling ability (Shaywitz & Shaywitz 2005). A longitudinal study on spelling development by Caravolas, Hulme & Snowling (2001) demonstrated that proficient spelling depends on both phoneme awareness and letter-sound knowledge. In this paper, we present evidence supporting this causal role through analyses of spelling errors within the setting of a spelling-level-match comparison between older children with spelling disabilities (SD) and younger normally progressing children. In addition, we compare the spelling-level-matched groups on language measures of phonological awareness, word recognition, phonological decoding, and orthographic coding. In the introduction, we review a theory suggesting shared mechanisms for spelling and reading, describe some recent studies with similar paradigms, and outline how the current approach tackles some of the issues facing the inconsistent findings in this line of research.

Romani, Olson, & Di Betta (2005) have argued that spelling and reading share the same orthographic and phonological representations. These representations are thought to be utilized within the framework of a dual-route model that posits a non-lexical and a lexical path. In the indirect phonological (non-lexical) route, phoneme-grapheme rules are used in the decoding

Correspondence concerning this article should be addressed to Angela Friend, Department of Psychology, University of Colorado, Boulder, CO 80309, USA.
Angela Friend, Department of Psychology; Richard K. Olson, Institute for Behavioral Genetics and Department of Psychology.

of written words into spoken words and vice versa, whereas in the lexical path, words are retrieved as whole units. A weakness in the phonological route may explain the pattern of deficits in decoding words in children with reading disabilities (RD). Under the premise of shared phonological and orthographic structures for reading and spelling, children with reading disabilities (RD) should not only demonstrate a phonological processing deficit in single word decoding but also a similar phonological processing deficit in spelling.

Several previous studies have investigated whether normally progressing children make the same kinds of spelling errors as older spelling-level-matched children with SD. Spelling-level-match comparisons typically involve pairing older spelling disabled (SD) children with younger normally progressing children on the number of words spelled correctly in a standardized spelling task. A few of these studies have found that normally progressing children tend to make errors that are more phonologically accurate than older children with SD. However, findings have been inconsistent.

Moats (1983) matched 27 4th to 8th grade children from a private school for children with dyslexia, who exhibited a minimum of a two-year lag in their reading achievement compared to their intellectual ability and a spelling grade level between 2.6 and 3.6 on the Stanford Dictated Spelling test, to 27 normally progressing children in second grade equated on their spelling performance. She conducted a detailed analysis of misspelled words for a variety of error types such as serial order confusions and phonological accuracy. Moats did not find significant group differences in the level of phonological accuracy, or in the number of serial order confusions. In her discussion, Moats stated that there was a moderate correlation between years of remediation training and phonological accuracy, and the majority of children with dyslexia in her study had between 7 months to 3 years of remedial instruction. Moreover, she stated that the children with dyslexia demonstrated the positive effects of phonics instruction. This suggests that the phonological remediation that the disabled group received may have eliminated or reduced apparent deficits in the phonological accuracy of spelling errors that would have been present without phonics intervention.

In contrast, Bruck and Treiman (1990) found a significant phonological deficit in the attempted spellings of children with SD in their study. They recruited their sample of 33 disabled spellers (mean age of 10 years and 2 months) from an assessment center for dyslexia, as well as 23 younger normally progressing children (mean age of 7 years and 5 months) from public schools. The groups were matched on a standardized spelling test. Bruck and Treiman’s (1990) data analysis demonstrated that the SD group made more initial consonant cluster errors (omission of a consonant) and thus made more errors that were considered non-phonetic, or not pronounceable in English.

Bourassa and Treiman (2003) conducted a study using a spelling-level matched design to compare oral and written spelling errors of children with SD (mean age 11 years and 1 month) from clinics specializing in tutoring for dyslexia, with that of younger children (mean age 7 years and 5 months) of normal reading and spelling ability. Their initial analysis consisted of a spelling sophistication composite based on a point system that gives more points to spellings where most or all of the phonemes in the target word are represented and higher scores are given to spellings that follow graphotactic conventions (i.e., use of legal letter sequences) and conventional spelling. They failed to find any significant group differences on their composite measure for either words or nonwords. The groups also performed similarly on graphotactic acceptability, wherein graphemic sequences are checked for whether they occur in English, and on the phonological skeleton, which measures how closely consonant and vowel sequences match target words. However, in a post hoc analysis of specific kinds of errors, they found a non-significant trend for higher graphotactic accuracy in the SD group, and some subtle but significant differences for specific types of errors. For example, older children with SD were

NIH-PA Author Manuscript

NIH-PA Author Manuscript

more likely to include a final vowel in words like supper, wrongly include a final “e” in words with a short vowel, and use single consonants for words requiring double consonants more often than younger normally progressing children. Bourassa and Treiman (2003) argued that these results were due to the SD children’s greater experience with patterns in print compared to the younger controls. Overall, however, this study does not support a phonological deficit in spelling and does not replicate the results of Bruck and Treiman (1990).

More recently, Cassar, Treiman, Moats, and Pollo (2005) compared 25 children identified with dyslexia through standard clinical practice from private schools and institutions that provide tutoring with that of 25 normally progressing children from public schools. The children with dyslexia had a mean age of 11 years and 7 months, whereas the normally progressing children had a mean age of 6 years and 8 months. The groups were matched on the spelling subtest of the WRAT-R (Jastak & Wilkinson, 1984). Analyses of phonological skills consisted of phoneme counting and nonword spelling, neither of which was found to demonstrate significant differences between groups. They also did not find any significant differences on their spelling choice task, which tested the children’s knowledge of legal letter patterns.

The present spelling-level-match study of phonological and graphotactic accuracy differs from the previous studies described here on several points. First, children with SD and the younger ability matched-controls came from the Colorado Learning Disabilities Research Center sample, which recruits solely from Colorado public schools. This is in contrast to the clinic samples used in some of the studies described earlier wherein phonological deficits may have been remediated. Many studies have demonstrated that tutoring in phonological awareness and phoneme-grapheme correspondences improves phonological accuracy in reading and spelling (cf. Wise, Ring, & Olson, 1999, 2000). It is possible that remediation of phonological deficits in the clinic-based samples may have eliminated or reduced differences between groups in their phonological spelling accuracy that would have been present without this intervention. However, with the exception of the study by Moats (1995), the studies described here do not report what type or amount of remediation that children with SD may have received.

In addition, the present sample is much larger than previous studies with 77 matched pairs of children and utilizes a within-subject design, which provides greater power for finding differences between the groups. Lastly, the present study included analyses of both spelling and reading measures, which allows for parallel comparisons of phonological and orthographic skills in spelling and in reading in the same sample.

Methods

Participants

One hundred and fifty-four children were selected for the current analyses from a larger, ongoing twin study conducted at the Colorado Learning Disabilities Research Center (CLDRC; DeFries, et al., 1997). Twins were recruited from school records from 27 Colorado school districts for participation in the CLDRC. They were selected for laboratory testing when at least one member of each twin pair showed some evidence of reading and/or spelling disability in their school records. In addition, a sample of twins, neither of whom had a school history for reading or spelling disability was also tested. Children were excluded from the study if a parental questionnaire revealed serious neurological problems, uncorrected vision or hearing deficits, serious social/emotional problems, or a first language other than English. Although the present study utilizes only data from twins, a large study comparing academic performance of adolescent twins and singletons found that performance was normally distributed with very similar means and standard deviations (8.02 vs. 8.02 and 1.05 vs. 1.06) when controlling for birth weight, gestational age at birth, age at test, parents age and education (Christensen, K. et al., 2006). The measure of academic achievement used consisted of teacher ratings and a general test given to Danish students in ninth grade. The test covered language, mathematics, hard and social sciences, and foreign languages.

To ensure independence of subjects, one twin of each pair was randomly selected within the school-history and no-school history groups. Then the groups were further selected based on their age-adjusted spelling performance relative to the no-school-history sample mean and standard deviation. Children with SD all had spelling scores at least -1 standard deviation below the no-school-history sample mean, and the children with no SD all had spelling scores above the one standard deviation cut off. The final selection of the two samples was done by finding a mean age and range for each group that allowed for the pairing of children with and without SD on raw spelling scores. This selection procedure yielded 77 matched pairs. The pair members with SD averaged 11.54 years of age (standard deviation = .31), and the pair members without SD averaged 8.56 years of age (standard deviation = .28). Most of the 77 pairs were exactly matched on their spelling score. The maximum within pair score difference was 3 points, and the group raw score mean was 20.88 (standard deviation = 2.61) for the SD group, and 20.82 (standard deviation = 2.58) for the control group.

Measures

Spelling: The Spelling subtest for children under the age of 12 from the Wide Range Achievement Test (WRAT-R; Jastak & Wilkinson, 1984) was administered to all children. The tester read words of increasing difficulty, first isolated and then in a sentence. Standardized pronunciations were provided in the testing manual. Administration continued until the child reached 10 misspelled words in a row or finished the list. The final score was the number of words spelled correctly, and this score was used for the matching of pairs. The appendix shows the words that were misspelled by at least one participant, and thus included in the error analysis.

Error Scoring: Spelling errors were scored by the first author, without knowledge of group membership, on percent graphotactic accuracy and percent phonological accuracy. Subjects' scores were computed by adding up individual word scores on phonological accuracy and graphotactic accuracy and dividing each total by the number of words spelled incorrectly, so that each child had an average graphotactic accuracy score, and an average phonological accuracy score.

Phonological Accuracy: Once pairs were matched on the number of words spelled correctly, a second score was computed for each child from their incorrect spellings. Scores were calculated by dividing the number of phonologically accurate phonemes represented in order within the attempted spelling over the total number of phonemes present in the conventional spelling of the word. Leeway was given for schwas, such that any vowel would be accepted, but not for short vowels or consonants that differ only in voicing. Words were also divided into syllables based on the syllable divisions presented in the Oxford Online Dictionary (accessed in 2004) in order to constrain phonemic interpretations. The following is a typical example of the phonological accuracy scoring method. First, the response [oder] for the word "order" was divided into two syllables [o – der]. The first syllable in the attempted spelling represents one of two phonemes correctly and the second syllable represents all three phonemes correctly, which yields a score of .80.

Although there is some disagreement as to where and whether syllable divisions are appropriate for English, we chose to add syllable divisions to the scoring of phonemes in order to have a phonological accuracy measure that would be sensitive to knowledge about phoneme-grapheme relationships and to how phonemes relate to each other within a word.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

If a child's spelling attempt was ambiguous as to where to make the syllable divisions, the word was always divided in such a way as to maximize the accuracy of the phonemes within each syllable. For example, the response [regsaball] for the word "reasonable" could be divided into [reg - s - a - ball]. In this case, the first syllable representation/reg/contains one accurate phoneme/r/because the letter "g" constrains the second grapheme to represent the short e sound rather than the correct long e sound. The second syllable contains no accurate phonemes, and the third syllable is a schwa so that any vowel would have been accepted. The fourth syllable also contains a schwa such that the syllable [ball] would be accepted to/bkl/would for a total phonological accuracy score .56. On the other hand, the same spelling could be divided [re - gs - a -ball] such that the first syllable/reI/would contain two correct phonemes instead of one for a total score .67. Therefore, the second spelling division was chosen because it yielded the highest score.

Additional letters were not explicitly counted against the total score, and a child could feasibly spell a word in such a way that it would be both phonologically inaccurate as a whole word because of additional letters and include all of the correct phonemes to get a perfect phonological accuracy score. However, we found that our method of counting the number of phonologically accurate phonemes within a syllable captured additional letters at syllable boundaries through their impact on the target phonemes, and that this method captured most instances of extra letters. The following is the only example where additional letters did not influence the representation of the phoneme and did not count against the total score. In this case, the child spelled the word "nature" as [matuchen], which was then divided into two syllables [ma – tuchen]. Although the child included an extra syllable, "tu" in their attempted spelling of "nature", their score was the same as if they had not included this additional syllable because it did not influence the/tʃ/phoneme in the second syllable.

A Guttman Split-half coefficient was computed by splitting the spelling words into an odd half and an even half, rather than first and second half because words increase in syllable and difficulty. The Guttman Split-half coefficient was .689 for the phonological accuracy score.

Graphotactic Accuracy: The graphotactic accuracy score was calculated in a similar manner as in Bourassa and Treiman (2003). Each attempted spelling was broken up into syllables based on the intended word, rather than by phonemes due to the variance in word length, and then each syllable was checked against the Oxford Online dictionary. In order to receive an accurate rating, the syllables must exist in current English words and in the same location (initial, medial, final). The score for each word was based on the number of graphotactically plausible syllables over the total number of syllables for a given word. The Guttman Split-half coefficient was .745 for the graphotactic accuracy score.

Reading: All participants were administered an extensive test battery for reading and related cognitive skills during two testing sessions at the University of Colorado. However, a small number of children have occasional missing scores and the degrees of freedom for some of the analyses reflect that. The subset of the entire battery of measures that was used in the current analyses is described here. Orthographic coding, phonological decoding, and phoneme awareness scores are composites comprised of the average percent correct scores from their component measures.

Word Recognition

Word Recognition: ***Timed Word Recognition Test (TWRT) (Olson, Forsberg, & Wise, 1994):*** The TWRT consists of words presented on a computer screen in order of increasing difficulty, as assessed in an independent sample. Oral naming responses are considered correct only when the correct pronunciation of the word is initiated within two seconds of stimulus onset. Testing continues through a list of 182 items until the participant fails to answer 10 of the last 20 items correctly within the time limit or the end of the list is reached. Scores are based on the last word read. Test-retest reliability is .93.

Orthographic Coding

Orthographic Coding: ***Homophone Choice (HOMPC) (Olson, Forsberg, & Wise, 1994):*** The Homophone Choice task required participants to select which of two homophones presented on the computer screen answered a question asked orally by the tester ("Which is a flower?" *rose* rows). There were 65 items. Reliability estimated from the correlation with Word-pseudohomophone Choice is .80.

Word-pseudohomophone Choice (ORTHPC) (Olson, Forsberg, & Wise, 1994): The Word-Pseudohomophone Choice task required participants to distinguish a real word from a nonword with the same pronunciation (rane *rain*). There were two sets of 40 items administered at the same time. The first half is easier than the second half. Split-half reliability is .93.

Orthographic Coding Composite: A composite measure was created by adding the percent correct homophone choice score and the word-pseudohomophone choice score and dividing it by two.

Phonological Decoding

Phonological Decoding: ***Phonological Choice (PHOPC) (Olson, Forsberg, & Wise, 1994):*** The Phonological Choice task consisted of 60 items requiring participants to select which of three nonwords would sound like a word (beal *bair* rabe). Reliability estimated from the correlation with oral nonword reading is .80.

Oral nonword reading (N1PC) (Olson, Forsberg, & Wise, 1994): The oral nonword reading task consisted of reading 45 one-syllable nonwords (ter, strale). Percent correct scores (plausible pronunciations) were calculated for the task. Test-retest reliability is .86.

Phonological Decoding Composite: A composite measure was creating by adding the percent correct phonological choice score and the percent correct oral nonword reading score and dividing it by two for each individual.

Phonological Awareness

Phonological Awareness: ***Phoneme Deletion (PDPC) (based on Bruce, 1964):*** The phoneme deletion task consisted of 68 trials presented via CD player in which the subject repeated a non- or word and was then asked to say it again, deleting a specified phoneme ("say prot – now say prot without the/r/"). Participants were given 2 s for repetitions and 4 s for deletion responses, as signaled by a warning tone on the CD, and a percent-correct score was calculated. Reliability estimated from the correlation with Phoneme Segmentation and Transposition is .80.

Phoneme Segmentation and Transposition task (WTDPIG) (Olson, Forsberg, & Wise, 1994) is a "pig-latin" game in which the participants take the first sound off the front of a word, put it at the end, and add the sound/ay/. For example, rope would become ope-ray. The test consists of nine practice and 45 test items with all words within the listening vocabulary of elementary school age children. There is no time constraint associated with the task.

Phonological Awareness Composite: A composite measure was creating by adding the percent correct phoneme deletion score and the percent correct phoneme segmentation and transposition score and dividing it by two for each individual.

Results

Spelling

The means and within-subject confidence intervals (Loftus & Masson, 1994) for phonological and graphotactic accuracy in spelling are presented in Figure 1. The results of the within-subject ANOVA demonstrated that neither the main effect of group F (1, 76) =.062, p = .804 or of measure F (1, 76) = .244, p =.623 was significant. However, the interaction of group by measure was significant F (1, 76) =8.72, p =.004. On average, within-subjects t-tests demonstrated that the mean performance for SD children was significantly below that of controls (79.7% vs. 82.0%) on the phonological accuracy score (PAspell) t (76) = −2.43, p = .018, but not significantly better (81.5% vs. 79.8%) on the graphotactic score (GRspell) t (76) = 1.28, p = . 21. Cohen's d measure of effect size for the phonological accuracy and graphotactic accuracy was .31 and .17 respectively.

Whereas the previous t-test of PAspell utilized all of the misspelled words, a second within-subject t-test of phonological accuracy using only the words that were misspelled by both members of each pair demonstrated a non-significant trend (80.2% vs. 82.0%) across groups. The trend toward greater phonological accuracy in controls was similar in direction to the results when using all of the misspelled words available t (76) = −1.77, p = .081. The smaller effect size (Cohen's d = .24 vs. .31) reflects the fact that the analysis was based on approximately half the number of words available.

Reading and Phonological awareness

The means and within-subject confidence intervals for the timed word recognition task (TWRT) phonological awareness composite (PAcomp) are presented in Figure 1. The SD group did not differ significantly from the controls (20.88 vs. 20.82) on the timed word recognition task t (75) = −.39, p =.70, which shows that for this sample, pair-matching the groups on raw spelling scores also results in a group-level match on word recognition (Cohen's *d = .02)*.

The mean and within-subject confidence intervals for phonological awareness are presented in Figure 1. This measure demonstrates a similar pattern to that of phonological decoding, (65.00 vs. 75.10) wherein the SD group mean is significantly below that of the control group F 50.62 (1, 73) = p < .001 (Cohen's d = .67).

We used measures of component skills in word reading, phonological decoding and orthographic coding, to run parallel analyses to those we ran for group differences in the phonological and graphotactic accuracy of spelling errors. The means and within-subject confidence intervals for orthographic coding (Orthcomp) and phonological decoding (PDcomp) are presented in Figure 1. The main effect of group F 30.20 (1, 68) = p < .001 and the main effect of measure F 13.84 (1, 68) = p < .001 were significant. The main effect of measure is not theoretically meaningful since there was no attempt to equate the difficulty of items for the two tasks, but it is important to note that the main effect of group is almost entirely due to the group difference in phonological decoding. Thus, there was a significant interaction of group by measure F 30.48 (1, 68) = p <. 001. On average, the SD group performed significantly below the controls (54.12 vs. 66.70) on the PDcomp t (73) = −.7.12, p < .001 (Cohen's d = .88), but not significantly different from the controls (65.41 vs. 66.15) on the Orthcomp t (70) = −.56, p = .58 (Cohen's d = .09), following the pattern found in the spelling error analysis.

To see if the within-group relations between variables are consistent with the between-group differences, correlations among spelling, reading and phonological awareness measures were calculated separately for each group (see Table 1). When controlling for WRAT-R spelling score, the PAspell measure in spelling correlated significantly with PAcomp for both the SD and control groups. In addition, PAcomp and PDcomp correlated significantly for both the SD and control groups. For the control group, the PAspell correlated with PAcomp $r = .270$, $p = .024$ but not with the PDcomp. Moreover, the non-significant correlation between PAspell and PDcomp in the control group is significantly different from the significant correlation between PAspell and PDcomp in the SD group (Fisher's $z = -2.66$, $p = .008$).

Discussion

Under the hypothesis of a general delay in the acquisition of spelling skills, older children with SD would be expected to demonstrate similar graphotactic and phonological accuracy in their spelling errors to younger controls matched on spelling performance. On the other hand, lower performance on phonological accuracy in spelling errors and similar performance on graphotactic accuracy would mirror typical findings in reading of a phonological deficit, and it would suggest shared orthographic and phonological representations for both reading and spelling. The analysis of variance of spelling errors showed that children with SD have significantly lower phonological accuracy than younger normally progressing controls matched on number of words spelled correctly. Moreover, our study found a significant group by measure interaction wherein children with SD had non-significantly higher graphotactic scores while scoring significantly lower on phonological accuracy. The pattern of results in spelling was similar to results in reading where there was a phonological decoding and awareness deficit along with no significant difference in orthographic coding. Therefore, the similar pattern of results for spelling and reading is consistent with the hypothesis of Romani et al. (2005) that there are shared orthographic and phonological representations in reading and spelling.

One issue that arises with our scoring method for phonological accuracy is that although the children are matched on overall spelling performance some of the words they misspell will differ within a matched pair. Words increase in difficulty and may differ in phoneme-grapheme correspondences as well as graphotactic sophistication. Therefore, we conducted a second t-test on only the subsection of words that were shared within pairs, such that they were matched both on the number of items they got correct and on the words misspelled. Although the difference was not significant, partly because the sample of words was reduced by approximately half, we found the same general trend and a similar effect size to the results utilizing all of the available words. This suggests that the significant difference we found in the first analyses is due to actual differences in phonological accuracy rather than differences in stimuli within pairs.

Although we found that children with SD are significantly less phonologically accurate in their spelling attempts than younger normally progressing children matched on spelling, the difference in accuracy is small and has not been found consistently in other studies. Studies investigating specific linguistic characteristics have typically found that children with SD do not differ significantly in the types of errors they make. Moreover, Cassar et al., (2005) found that experienced teachers were not able to differentiate the spelling attempts of older children with SD from that of younger controls. These results suggest that the types of errors that children with SD make are similar to that of younger controls and that the slightly smaller magnitude of phonological accuracy found in the present study would not be applicable as a diagnostic in the classroom.

Studies utilizing a spelling-level-match design have not yielded consistent results in support of a phonological deficit in spelling. The inconsistent findings in the literature may be due to several factors. For example, the assessment of phonological and graphotactic accuracy in spelling errors varies across studies. Another likely source of differences across some of the other studies and the present study may be related to power. We have a much larger sample and utilized within subject analyses to test our comparisons, which likely improved our ability to find small differences in phonological accuracy. Lastly, there were methodological differences in the selection of subjects that may have resulted in different amounts of remedial instruction in phoneme-grapheme correspondences. Moats (1983) found that exposure to the Orton-Gillingham program, which focuses on phonetic instruction, could account for a moderate proportion of the variance in phonetic spelling in the older disabled group. Furthermore, she stated that there was a significant correlation between years of remediation training and phonological accuracy in her study. Cassar et al. (2005) and Bourassa and Treiman (2003) also recruited their samples of children with SD from clinics and failed to find significant differences, whereas significant differences were found by Bruck and Treiman (1990), who recruited children from an assessment center that did not provide remediation, and by the current study, which recruited from public schools. However, although we cannot be certain that the children in the clinic-based studies had more specific training in phonological awareness and grapheme-phoneme correspondences than our school-based sample, it does offer one possible account of why a phonological deficit in spelling has not been consistently found in the literature.

Previous spelling-level-match studies of phonological and graphotactic accuracy in spelling have not included parallel assessments of phonological awareness in language or of phonological decoding and orthographic coding component skills in word reading. Therefore, we can not tell if the clinic-based spelling-disabled samples in previous studies might have had their phonological processing deficits remediated in both spelling and reading. We have already noted that on average in our school-based spelling-disabled group, there were parallel patterns of phonological deficit in reading and spelling, and in phonological awareness. However, we found considerable within-group variation in component spelling, reading, and language skills, even after controlling for overall accuracy on the WRAT-R spelling test. Our hypothesis of a linkage between these skills in reading and spelling, based on our group-means comparisons, would be strengthened if there were significant within-group correlations showing a link between phonological processes in reading and spelling.

The correlations in Table 1 for the spelling-disabled group below the diagonal showed this linkage for within-group individual differences quite clearly. Phonological awareness was significantly correlated with phonological accuracy in spelling ($r = .47$) and phonological decoding ($r = .64$), but not significantly with graphotactic accuracy in spelling ($r = .11$) or reading ($r = -.04$), after controlling for overall WRAT-R spelling score. Phonological accuracy in spelling was also significantly correlated with phonological decoding accuracy in reading ($r = .42$), and it was not significantly correlated with orthographic coding accuracy in reading ($r = .003$). Thus, the individual differences within our spelling-disabled group provide strong support for a linkage between phonological skills in reading and spelling.

The correlations with phonological awareness within the younger normally progressing group were also significant with phonological spelling accuracy ($r = .27$) and phonological decoding ($r = .40$), suggesting that even among normally progressing spellers and readers, there are reliable individual differences related to phonological skills (c.f., Bryant and Impey, 1986). These correlations were non-significantly lower than in the spelling disabled group, and the normal group's correlation between phonological spelling accuracy and phonological decoding in reading ($r = -.01$) was significantly lower than the correlation within the spelling disabled group. This pattern of generally lower correlations within the normally progressing group may be partly due to some of the children in our school-based spelling-disabled sample having received systematic remediation for their phonological processing deficits in reading and spelling that led to that group's stronger correlations among phonological awareness and phonological skills in spelling and reading. Unfortunately, we do not have data on phonologically based remedial interventions for the children in the spelling disabled group, but experimental studies have shown that such interventions can have strong and highly correlated influences on phonological awareness and on phonological skills in reading and spelling (cf. Wise et al., 2000).

We have shown that there is a linkage between phonological skills in reading and spelling across groups and that the SD group demonstrates a phonological deficit in both reading and spelling. However, the effect size for the phonological accuracy deficit in spelling (.34) was much lower in magnitude than either the phonological awareness deficit (.72) or the phonological decoding deficit in reading (.84). There are several possible reasons that may have led to the smaller effect size for spelling. One reason may simply be because we have a lower reliability for the phonological spelling accuracy measure than for the measures that make up the phonological awareness and decoding composite scores, which introduces greater error variance around the phonological spelling accuracy estimates and reduces the effect size. Another possible reason may be due to the asymmetry in the conversion of phonemes to graphemes and vice versa inherent to English. The asymmetry leads to differences in consistency in reading and spelling, such that spelling is always the less consistent of the two (Kessler & Treiman, 2001). For example, there are over 170 graphemes ranging from single letters to groups of letters such as “gh” to represent approximately 42 different phonemes in English (Moats, 1995). The spelling accuracy measure accepts all plausible graphemes within each syllable, which may lead to less variance available for capturing phonological skill differences. Thus, the difference in magnitude of effect sizes between phonological deficits in reading and spelling may be due to a greater availability of plausible responses in spelling, as well as the lower reliability for our measure of phonological accuracy in spelling errors on the WRAT-R..

In summary, older children with SD demonstrated a similar pattern of significant group deficits in both the phonological accuracy of their spelling errors, their phonological decoding accuracy in reading, and in their phoneme awareness, compared to a group of younger normally progressing spellers matched on number of words spelled correctly on the WRAT-R. Moreover, the linkage between the SD group’s deficits in reading and spelling was complemented by our finding of significant within-group correlations between individual SD children’s phonological awareness, their phonological accuracy in spelling, and their phonological decoding accuracy in reading. We suggested that some previous failures to find a significant phonological spelling accuracy group deficit in children with SD may have been due to their phonological remediation in those clinic-based samples, but this hypothesis should be tested experimentally for SD children with and without such remediation, and with parallel measures in spelling and reading.

Acknowledgments

This research was supported by grants from the National Institute of Child Health and Human Development, HD27802 and Training Grant HD-007289.

We are especially grateful to Lise Menn for her expertise in linguistics during the early phases of creating the error scoring methods and to Jacqueline Hulslander for statistical support.

References

Bourassa D, Treiman R. Spelling in Children with Dyslexia: Analyses from the Treiman-Bourassa Early Spelling Test. Scientific Studies of Reading 2003;7:309–333.

Bruce DJ. The analysis of word sounds by young children. British Journal of Educational Psychology 1964;34(2):158–170.

Bruck M, Treiman R. Phonological Awareness and Spelling in Normal Children and Dyslexics: The Case of Initial consonant Clusters. Journal of Memory & Language 1990;45(4):751–775.

Bryant PE, Goswani UC. Strengths and weaknesses of the reading level reading level design: A comment on Blackman, Mamen, and Ferguson. Psychological Bulletin 1986;100:101–103.

Bryant PE, Impey L. The similarities between normal readers and developmental and acquired dyslexics. Cognition 1986;24:121–137. [PubMed: 3791918]

Caravolas M, Hulme C, Snowling MJ. The foundations of spelling ability: evidence from a 3-year longitudinal study. Journal of Memory and Language 2001;45:751–774.

Cassar M, Treiman R, Moats L, Pollo TC, Kessler B. How do the spellings of children with dyslexia compare with those of nondyslexic children? Reading and Writing 2005;18:27–49.

Christensen K, Petersen I, Skytthe A, Herskind AM, McGue M, Bingley P. Comparison of academic performance of twins and singletons in adolescence: follow-up study. British Medical Journal 2006;333:1095–1097. [PubMed: 17012267]

DeFries JC, Filipek PA, Fulker DW, Olson RK, Pennington BF, Smith SD, et al. Colorado Learning Disabilities. Research Center. Learning Disability Quarterly 1997;8:7–19.

Jastak, S.; Wilkinson, GS. Wide Range Achievement Test–Revised (WRAT-R). Jastak Assessment Systems; 1984.

Kessler B, Treiman R. Relationships between sounds and letters in English monosyllables. Journal of Memory and Language 2001;44:592–617.

Loftus GR, Masson ME. Using confidence intervals in within-subject designs. Psychonomic Bulletin & Review 1994;1(4):476–490.

Moats LC. A comparison of the Spelling Errors of Older Dyslexic and Second Grade Normal Children. Annals of Dyslexia 1983;33:121–140.

Moats, LC. Spelling development disability and instruction. Baltimore, MD: York Press; 1995.

Olson, R.; Forsberg, H.; Wise, B. Measurement of word recognition, orthographic, and phonological skills. In: Reid, LG., editor. Frames of reference for the assessment of learning disabilities: New views on measurement issues. Baltimore: Paul H. Brookes Publishing; 1994. p. 243-277.p. 243-277.

Oxford English Dictionary online. Oxford University Press 2005. All rights reserved. Retrieved March 2–28, 2005 from http://dictionary.oed.com/entrance.dtl

Rack JP, Snowling MJ, Olson RK. The nonword reading deficit in developmental dyslexia: A review. Reading Research Quarterly 1992;27:28–53.

Romani, C.; Olson, A.; Di Betta, AM. Spelling Disorders. In: Snowling, MJ.; Hulme, C., editors. The science of reading: A handbook. Malden, MA: Blackwell publishing Ltd; 2005. p. 431-448.

Shaywitz SE, Shaywitz BA. Dyslexia (specific reading disability). Biological Psychiatry 2005;57(11): 1301–1309. [PubMed: 15950002]

Vellutino FR, Fletcher JM, Snowling MJ, Scanlon DM. Specific reading disability (dyslexia): what have we learned in the past four decades? Journal of Child Psychology and Psychiatry 2004;45(1):2–40. [PubMed: 14959801]

Wise BW, Ring J, Olson RK. Training phonological awareness with and without attention to articulation. Journal of Experimental Child Psychology 1999;72:271–304. [PubMed: 10074381]

Wise BW, Ring J, Olson RK. Individual differences in gains from computer-assisted remedial reading with more emphasis on phonological analysis or accurate reading in context. Journal of Experimental Child Psychology 2000;77:197–235. [PubMed: 11023657]

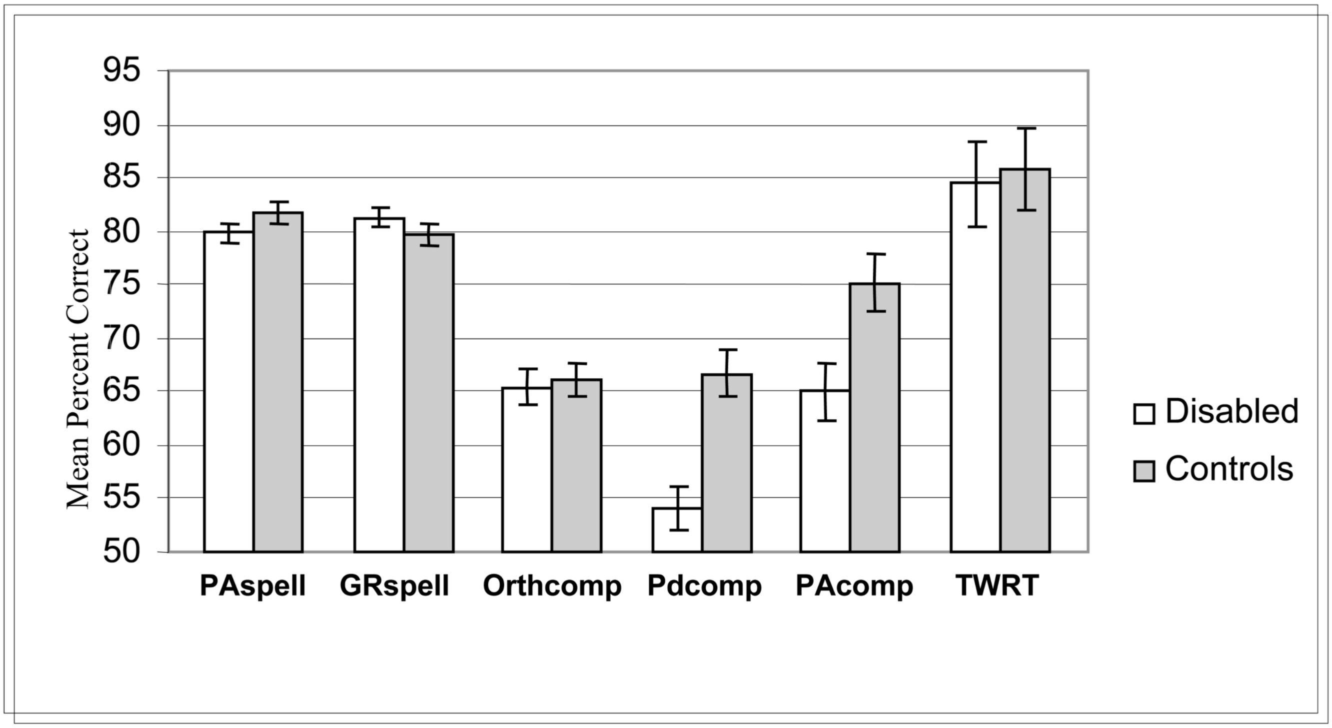

Figure 1.
Means and standard errors for PASPEL (phonological accuracy in spelling errors), GRSPEL (graphotactic accuracy in spelling errors), PDcomp (phonological decoding composite), Orthcomp (orthographic coding), PAcomp (phonological awareness composite), and TWRT (time limited word recognition test).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Partial correlations between spelling, reading, and phonological awareness measures for children with SD and controls when controlling WRAT-R spelling performance

Measures	PAspell	Orthspell	PDcomp	Orthcomp	PAcomp
PAspell	—	.361*	−.014	−.147	.270*
Orthspell	.421*	—	.018	−.152	.169
PDcomp	.417*	.099	—	.144	.400*
Orthcomp	.003	−.047	.110	—	−.104
PAcomp	.467*	.105	.641*	−.044	—

Note. PAspell = spelling accuracy measure, Orthspell = graphotactic accuracy measure in spelling, PDcomp = phonological decoding composite, Orthcomp = orthographic coding, and PAcomp = phonological awareness composite. Controls are presented above the diagonal. Asterisk indicates significance.

Appendix

List of misspelled words, syllable separations, and pronunciations.

1. make	meɪk	2. cut	kʌt
3. cook	kʊk	4. light	laɪt
5. must	mʌst	6. dress	dres
7. reach	ritʃ	8. or-der	Ɔrdər
9. watch	wɒtʃ	10. en-ter	ɛntər
11. grown	groʊn	12. na-ture	neɪtʃər
12. ex-plain	ɪkspleɪn	13. edge	ɛdƷ
14. kitch-en	kɪtʃən	15. sur-prise	sərpraɪz
16. ad-vice	ædvaɪz	17. pur-chase	Pɜrtʃəs
18. brief	brif	19. suc-cess	səksɛs
20. rea-son-a-ble	rizənəbəl	20. i-mag-i-na-ry	ɪmædƷənɛri
21. occ-u-py	ɒkyəpaɪ	22. char-ac-ter	kærɪktər
23. so-ci-e-ty	səsaɪɪti	24. o-ffic-ial	əfɪʃəl
25. rec-og-nize	rɛkəgnaɪz	26. fa-mil-iar	fəmɪlyər
27. co-mmiss-ion	kəmɪʃən	28. ben-e-fic-ial	bɛnəfɪʃəl
39. a-ppro-pri-a-tion	əproʊprieɪʃən	30. en-thu-si-as-m	ɛn θuziæzəm
31. crit-i-cize	krɪtəsaɪz	32. pre-ju-dice	prɛdƷədɪs
33. be-llig-e-rent	bəlɪdƷərənt	34. o-ccurr-ence	əkƷrəns