Available online at www.sciencedirect.com

SCIENCE DIRECT®

J. Experimental Child Psychology 87 (2004) 125–154

Journal of Experimental Child Psychology

www.elsevier.com/locate/jecp

Variation among developmental dyslexics: Evidence from a printed-word-learning task

Caroline E. Bailey,[a] Franklin R. Manis,[a,*] William C. Pedersen,[b] and Mark S. Seidenberg[c]

[a] *Psychology Department, University of Southern California, Los Angeles, CA 90089-1061, USA*
[b] *California State University, Long Beach, CA 90840, USA*
[c] *University of Wisconsin at Madison, Madison, WI 53706, USA*

Received 25 March 2003; revised 21 October 2003

Abstract

A word-learning task was used to investigate variation among developmental dyslexics classified as phonological and surface dyslexics. Dyslexic children and chronological age (CA)- and reading level (RL)-matched normal readers were taught to pronounce novel nonsense words such as *veep*. Words were assigned either a regular (e.g., "veep") or an irregular (e.g., "vip") pronunciation. Phonological dyslexics learned both regular and exception words more slowly than the normal readers and, unlike the other groups, did not show a regular-word advantage. Surface dyslexics also learned regular and exception words more slowly than the CA group, consistent with a specific problem in mastering arbitrary item-specific pronunciations, but their performance resembled that of the RL group. The results parallel earlier findings from Manis,Seidenberg, Doi, McBride-Chang, & Petersen [*Cognition 58* (1996) 157–195] indicating that surface dyslexics and phonological dyslexics have a different profile of reading deficits, with surface dyslexics resembling younger normal readers and phonological dyslexics showing a specific phonological deficit. Models of reading and reading disability need to account for the heterogeneity in reading processes among dyslexic children.

Keywords: Reading; Dyslexia; Subgroups; Phonological skill; Word identification; Models of reading

---

[*] Corresponding author. Fax: 1-213-746-9082.
E-mail address: manis@usc.edu (F.R. Manis).

0022-0965/$ - see front matter
doi:10.1016/j.jecp.2003.10.004

Introduction

There is a broad consensus among reading researchers that phonological deficits are a core problem in developmental dyslexia (Bradley & Bryant, 1983; Lyon, 1995; Shankweiler et al., 1995; Share, 1995; Stanovich, 1988; Wagner & Torgesen, 1987). However, reading is a complex process, and it seems likely that the developing reading system may fail in multiple ways (Pennington, 1999; Seidenberg, 1993). This line of thinking has led to many attempts in the past 3 decades to classify reading-disabled children into subtypes (e.g., Boder, 1973; Doehring, Trites, Patel, & Fiedorowicz, 1981; Lyon & Watson, 1981; Morris, Blashfield, & Satz, 1981). Much of the early work on subtypes was not based on explicit, well-developed models of the reading process and hence did not converge on distinct, replicable subtypes linked to hypotheses about causes, with the exception of Boder (1973). There has been progress in describing the range of cognitive deficits underlying reading disability subtypes. In a cluster analytic study of a large representative sample of poor readers, oral language and phonological processing deficits were found to be the most common concomitants of reading difficulties (Morris et al., 1998).

However, it is important to explore patterns of reading disability from the perspective of precise models of reading (Coltheart, Curtis, Atkins, & Haller, 1993; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; Share & Stanovich, 1995). Some models make predictions about patterns of individual differences in reading development, which has led to a recent increase in studies of dyslexic subtypes, both as a means of testing the theories and as an attempt to clarify the bases of reading impairments. The purpose of this study was to test theoretical explanations for two subtypes of reading difficulty, phonological dyslexia and surface dyslexia.

Two patterns corresponding to the phonological/surface distinction in the literature have figured prominently in recent studies of variability among dyslexic children (Castles & Coltheart, 1993; Castles, Datta, Gayan, & Olson, 1999; Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Manis et al., 1999; Murphy & Pollatsek, 1994; Sprenger-Charolles, Cole, Lacert, & Serniclaes, 2000; Stanovich, Siegel, & Gottardo, 1997). Nonwords (such as *mape* and *veet*) and exception words (such as *have*, *people*, and *yacht*) have been utilized to define the groups. Phonological dyslexics have difficulty pronouncing both nonwords and exceptions, but the degree of impairment is greater on nonwords. In addition, they show associated deficits in phoneme analysis and segmentation tasks. The "surface" or "delay" pattern of performance involves difficulty in pronouncing both nonwords and exceptions, but the degree of impairment is greater on exceptions. Most dyslexics have considerable difficulty reading both nonwords and exceptions. Researchers agree that these profiles of reading failure can be identified in children, although there is disagreement about whether the profiles are qualitatively distinct, or graded, and about the underlying causes of the reading profiles. Three different theoretical frameworks have been offered to account for the subgroup patterns, a dual-route framework, a connectionist framework, and a phonological core deficit framework.

According to dual-route models of reading (Coltheart, 1978; Coltheart et al., 1993; Frith, 1985; Patterson & Morton, 1985) skilled readers have two distinct mechanisms for pronouncing printed words, a lexical procedure, which involves accessing representations of whole words, and a sublexical procedure, which involves using spelling-to-sound correspondence rules to pronounce printed words. Castles and Coltheart (1993) proposed that these two mechanisms are distinct in developing readers as well as skilled readers, and hence there should be two main patterns of developmental dyslexia. Phonological dyslexia results from difficulties in using the sublexical procedure and surface dyslexia from a complementary difficulty with the lexical procedure. Several case studies exhibiting pure phonological dyslexia or surface dyslexia profiles have been reported (e.g., Coltheart, Masterson, Byng, Prior, & Riddoch, 1983; Goulandris & Snowling, 1991; Hanley, Hastie, & Kay, 1992; Temple & Marshall, 1983).

Castles and Coltheart (1993) assessed the frequency of phonological dyslexia and surface dyslexia profiles in a sample of 53 dyslexic children, ages 7 to 14 years. They identified 10 pure surface dyslexics who had nonword reading within the range for age-matched normal readers, but low exception word reading, and 8 pure phonological dyslexics, who showed the opposite profile. An additional 27 children were poor on both tasks, and hence could be classified as "mixed" cases, but showed a discrepancy between exception word and nonsense word reading. Most were more impaired on nonsense word reading (21 of 27) and therefore classified as phonological dyslexics. The remainder of the cases did not show a large enough discrepancy to meet subgroup guidelines adopted by the authors. Castles and Coltheart (1993) concluded that phonological dyslexics and surface dyslexics were fairly common (45 of 53 cases) among typical dyslexic samples and that they could be interpreted as different levels of development of the lexical and sublexical mechanisms.

Subsequent empirical work has replicated Castles and Coltheart's (1993) findings that phonological dyslexics and surface dyslexics occur commonly in dyslexic samples, but offered alternative interpretations. Manis et al. (1996) classified a sample of 51 dyslexic children (ages 10–15 years) as phonological dyslexic or surface dyslexic using relative performance on exception words and nonsense words. They found relatively few pure cases (5 per subgroup) (see also Murphy & Pollatsek, 1994). Most dyslexics had deficits on both nonwords and exception words, but a substantial number differed in the degree of deficiency, as Castles and Coltheart (1993) had observed. Manis et al. (1996) showed that the subgroups differed on independent measures of phonological and orthographic skills in predicted ways. The phonological dyslexia group ($n = 18$) performed more poorly than the surface dyslexia group ($n = 17$) on a task that required segmental analysis of spoken nonwords, whereas the surface dyslexia group performed more poorly than the phonological dyslexia group on an orthographic judgment task in which they had to decide which of two phonemically equivalent spellings of a word (e.g., *rane* vs *rain*) was a correct spelling of a word (e.g., Olson, Wise, Connors, Rack, & Fulker, 1989). Phonological dyslexics were also more likely than surface dyslexics to make exception word pronunciation errors that involved visual approximations to the target word, but less likely to make phonologically appropriate errors. Phonological dyslexics performed more poorly than age-matched normal readers on all reading and phoneme awareness tasks and more poorly than reading-level- (RL) matched normal readers on all phonological but not orthographic tasks. Surface dyslexics were strikingly similar to the RL comparison group on all measures, including error patterns.

Manis et al. (1996) hypothesized that phonological dyslexics fit a pattern of developmental deficit in phonological processing, whereas surface dyslexics fit a pattern of developmental delay in overall reading (based on the comparisons to the RL group). It is important to point out that surface dyslexics did not have general cognitive delays, as indicated by performance in the average range on IQ subtests. Bryant and Impey (1986) made a similar argument with regard to comparisons of surface dyslexia cases to younger normal readers. Based on connectionist models developed by Seidenberg and McClelland (1989) and Plaut et al. (1996), Manis et al. (1996) proposed that the developmentally deviant profile of phonological dyslexics was caused by impaired phonological representations. In contrast, the developmentally delayed reading pattern in surface dyslexics could be attributable to orthographic deficits or to a rate of learning or resource limitation, both of which have been shown in connectionist simulations to slow down the process of forming orthographic–phonological connections and to produce the profile of exception word lower than nonsense word performance.

In an important extension of the modeling work, Harm and Seidenberg (1999) demonstrated that a connectionist system learning to read with faulty phonological representations simulated the performance of phonological dyslexic children very closely. A number of anomalies (e.g., removing phonological units, adding Gaussian noise to the phonological activations) impaired the model's capacity to represent phonological codes. Phonological deficits of varying degrees always resulted in poor nonsense word performance relative to exception words. Severe phonological impairment resulted in a mixed profile, with deficits in both exception and nonsense words, and mild phonological impairment resulted in a pure phonological dyslexic profile (normal on exception words and below normal on nonsense words). Assuming that most dyslexics have phonological deficits, this approach explains why the majority of dyslexics show a mixed profile (deficient on both nonwords and exceptions). An important point is that the model did not address the fundamental basis for the phonological impairment (i.e., whether it is internal to the phonological system or derives from a more peripheral sensory processing deficit).

Harm and Seidenberg (1999) suggested that the delay characteristic of surface dyslexia could arise from several causes. Leaving phonology intact and reducing the number of hidden units (an intervening layer of units between orthographic and phonological representations) produced something very much like the surface dyslexia profile. A small reduction in the number of hidden units mainly affected the model's capacity to learn the pronunciations of words with atypical spelling–sound correspondences ("exception" or "strange" words such as *have*, *friend*, and *aisle*). With a more severe reduction of hidden units, the capacity of the model to encode more consistent mappings was also impaired, affecting generalization to novel regular words. Other types of "endogenous" deficits, such as inefficiency in the learning algorithm that caused the model to benefit less than normally from individual experiences with words, also produced the surface dyslexia/delayed pattern. However, exogenous deficits could also produce the surface dyslexia pattern, such as a lack of reading experience, or a curriculum overemphasizing regularities in print–sound correspondence. The simulations explain why exception words are most vulnerable in surface dyslexia and why nonword generalization is often somewhat impaired as well (i.e., mixed cases are more common than pure cases). The simulations also explain the profile of developmental reading delay. The impairments that created a developmental delay had more impact on learning exceptions than regularities, which parallels the pattern of relative difficulty observed in normal acquisition (e.g., Backman, Mamen, & Ferguson, 1984).

A third account of phonological dyslexia and surface dyslexia was provided by Stanovich et al. (1997). They gave multiple measures of reading to 68 third-grade dyslexics and appropriate groups of chronological age (CA)- and RL-matched normal readers. As in the Manis et al. (1996) study, surface dyslexics performed much like younger RL-matched normal readers across a wide spectrum of reading-related tasks, whereas phonological dyslexics presented a developmentally deviant profile (poor performance for reading level on phonological tasks and equal or superior performance on orthographic tasks). Stanovich et al. (1997) differed from both the dual-route and the connectionist accounts insofar as they argued that both the surface dyslexia and the phonological dyslexia patterns derive from a core phonological deficit. They proposed that surface dyslexics have a mild phonological impairment coupled with inadequate reading experience. Together these factors create a developmental delay. Their view also predicts that surface dyslexics should exhibit deficits on measures of print exposure, such as the Title Recognition Test (Cunningham & Stanovich, 1990; Stanovich & West, 1989).

The Stanovich et al. proposal raises two issues. First, do surface dyslexics have a mild phonological impairment that is the proximal cause of their reading problems? In the Manis et al. (1996) and Stanovich et al. (1997) studies, the surface dyslexics' performance on measures related to phonological knowledge was like that of younger normal readers. Stanovich et al. (1997) emphasized the fact that they performed less accurately than same-age good readers, indicative of a phonological impairment, but one that is milder than in phonological dyslexia. However, another way of viewing the data is that the surface dyslexics' phonological skills were normal given their level of reading achievement. This is important to consider because it is known that phonological awareness is shaped by experience in reading alphabetic orthographies (Ehri, Wilce, & Taylor, 1987; Morais, Cary, Alegria, & Bertelson, 1979; Perfetti, Beck, Bell, & Hughes, 1987). Hence, it could be argued that surface dyslexics have a normal phonological system, but perform as well on measures of phonological skill as would be expected from their progress in learning to read.

The second issue is whether surface dyslexics are in fact experientially deprived. Stanovich et al. (1997) did not test this prediction directly. Manis et al. (1999) investigated this issue in a sample of 72 dyslexic children tested longitudinally in the third and fourth grades, but failed to observe differences in print exposure (book title recognition task) between subgroups of phonological dyslexics and surface dyslexics. However, phonological dyslexia was found to be more stable over a 1-year period than surface dyslexia, a finding that is consistent with a stronger environmental influence on surface dyslexia. Other interesting observations from this study were that the mixed profile was by far the most common, and the surface dyslexia group showed a mild deficit in phoneme awareness relative to both reading- and age-level-matched normal readers. Whether the phoneme awareness deficit was a result or a cause of reading failure was unclear.

The most thorough study of experiential (and genetic) factors to date is that of Castles et al. (1999), who investigated dyslexic subgroups using a large dyslexic twin data base (e.g., Olson et al., 1989). Children ages 8 to 18 years were classified as phonological dyslexics or surface dyslexics, based on a discrepancy between exception word and nonsense word reading. The sample sizes were quite large (322 per subgroup). They found that word reading deficits were significantly heritable for both subtypes, but the genetic contribution to the group reading deficit was much greater in the phonological dyslexics (67% of the variance) than in the surface dyslexics (31% of the variance). In contrast, the shared environmental contribution to the group word reading deficit was 63% for the surface dyslexics. The subgroups differed significantly in the appropriate directions on measures of orthographic and phonological knowledge, phoneme deletion, regularization errors in word pronunciation (e.g., pronouncing *have* to rhyme with *gave*), lexicalization errors in word pronunciation (e.g., saying "sad" for *said*), and print exposure (book and magazine title recognition).

Castles et al. (1999) concluded that genetic deficits in phonological processing were prominent in phonological dyslexia, whereas environmental factors (such as print exposure and reading instruction) make a greater contribution to surface dyslexia. The differences in title recognition provide some support for the hypothesis of Stanovich et al. (1997) that print exposure plays a role in surface dyslexia, but the magnitude of the differences between groups was small (less than a third of a standard deviation). Hence, other shared environmental factors not measured in the study might be involved in surface dyslexia.

To summarize, the behavioral differences between phonological dyslexics and surface dyslexics are fairly well established but there is disagreement about their bases. Existing accounts overlap in many ways but also differ in detail. The dual-route and connectionist models entail radically different assumptions about how lexical knowledge is represented, learned, and used in performance and about the bases of the phonological dyslexia and surface dyslexia patterns. The account of Stanovich et al. shares some assumptions with the dual-route approach, but differs from the other two models insofar as it suggests that phonological dyslexia and surface dyslexia derive from different extents of phonological impairment and impoverished experience.

Most previous studies have focused on performance in reading words and nonwords aloud. The present study addressed questions about the bases of differences among dyslexics in a different way, focusing on the learning of new pronunciations. The study built on previous work by Castles and Holmes (1996). They taught pure phonological dyslexics and surface dyslexics ($n = 8$ per subgroup, 8–13 years of age) the pronunciations of 20 nonsense words that were assigned exceptional pronunciations. Half of the nonsense words had commonly occurring spelling patterns (e.g., *weaf* pronounced as rhyming with "deaf"); these were termed exceptions. The other half had unusual orthography (e.g., *macht* pronounced as rhyming with "yacht"); these were termed "strange" words (Seidenberg, Waters, Barnes, & Tanenhaus, 1984). The children learned the items over the course of four 15-min training sessions with 3 or 4 days between sessions. They were tested using a reading-aloud task as well as an orthographic choice task (e.g., choose the correct spelling from homophonous alternatives such as *mot* and *macht*). An important advantage of this paradigm is that one aspect of print exposure is equated: the number of positive feedback trials with each word. Thus, any group differences in learning rate would be attributable to differences in word reading processes rather than amount of practice.

Castles and Holmes (1996) found that surface dyslexics were less accurate than phonological dyslexics at reading the strange items but not the exceptions. Surface dyslexics also performed less accurately than phonological dyslexics across both word types on the orthographic choice task. The results indicated that surface dyslexics were impaired in learning exception words even when the number of exposures to the words was experimentally controlled, suggesting that factors beyond sheer amount of practice in reading words are entailed in surface dyslexia. Castles and Holmes (1996) interpreted the data as evidence for deficient functioning in two qualitatively distinct mechanisms, either the lexical or the nonlexical mechanism. However, a connectionist model could account for the same data as a difference in spelling–sound consistency. Strange items and exceptions are harder for surface dyslexics due to systemic deficits (such as reduced hidden units or poor orthographic codes) affecting the rate of item-specific learning.

Although the paradigm of Castles and Holmes (1996) provides an interesting way to evaluate individual differences among dyslexics, there were four key limitations. First, the amount of practice at reading the novel words was equated across groups and word types, but because most of the exception and strange word pronunciations were based on relatively common words (e.g., *couch*, *shoe*, *month*, *prove*, *deaf*), it is possible that subgroup differences might depend on differences in prior experience with words that resemble the novel words. If print exposure was indeed lower in the surface dyslexia group, they would be less likely to have encountered the exception and strange words on which the stimuli were based. In a connectionist model, prior exposure to printed words containing these exceptional spelling patterns and pronunciations would have an effect on the ability to learn new instances of the same spelling patterns. Lower frequency base words should be used, or the base words should be given both regular and exception word pronunciations to control for familiarity with the item as an orthographic unit. Second, the study did not compare dyslexics' performance to that of younger normal readers and hence cannot shed light on the delay vs deficit distinction. Including younger RL-matched readers is also valuable because it provides a type of control for differences in overall knowledge of printed words and reduces the likelihood that the patterns of performance of the subgroups are solely due to prior experience in reading. Third, the absence of CA controls means we cannot determine the extent of deficiency in word learning among the two groups. Finally, performance on novel words with regular pronunciations was not examined, so it cannot be determined whether the surface dyslexia children's deficit was specific to the learning of strange or exception words.

In the current study, a similar word learning task was administered to fourth- and fifth-grade poor readers classified as phonological dyslexics or surface dyslexics, as well as to age- and reading-level-matched normal readers. Children were required to learn the pronunciations of 22 printed nonsense words that were said to be the names of "space creatures." Half of these novel words were assigned a regular and half an exceptional pronunciation for each child. Each word was used as its own control by administering counterbalanced lists to half of the participants in each group. For example, *zide* was pronounced /zayd/ (phonologically regular) for half of the participants and /zId/ (exception word) for the other half. Hence, the design controls for differences in visual and orthographic processing of the letter strings between the regular and the exception word conditions, a control that is missing in comparisons of the ability to read regular and exception words in the literature. We obtained independent measures of orthographic knowledge, phoneme analysis, and print exposure to validate the subgroups and to test alternative hypotheses about the source of differences between phonological dyslexics and surface dyslexics.

The theories predict distinctly different patterns of results across all of the measures. The dual-route model predicts that phonological dyslexics should have a problem learning the rule-governed items, but not the exceptions, given an abnormal sublexical and a normal lexical mechanism. Surface dyslexics should have a problem learning the exceptions but not the regular words, because the latter can be read successfully with an intact sublexical mechanism. Phonological dyslexics should show a smaller than normal regularity effect, and surface dyslexics should show a normal or even larger than normal regularity effect (owing to the poor functioning of their lexical system).

Stanovich et al. (1997) proposed that phonological dyslexia and surface dyslexia cases differ in the degree of phonological impairment, but that surface dyslexics have lower print exposure. Although print exposure, as typically measured, includes more than just number of exposures to printed words, the key aspect in surface dyslexia is assumed to be experience with specific words. Because the number of exposures to the novel words is controlled in the word-learning task, and the subgroups' overall knowledge of printed words is equated for the subgroups and the RL group, the major differences between dyslexics and normal readers on this task should involve phonological processing. Surface dyslexics should perform more accurately overall than phonological dyslexics, to the extent that any novel word (even exceptions) contains phonologically regular elements. Due to the phonological deficit, both subgroups should show smaller than normal regularity effects in comparison to both CAs and RLs. This approach also predicts that surface dyslexics should have lower print exposure than phonological dyslexics on an independent measure of this construct.

Unlike the proposal of Stanovich et al. (1997), the connectionist model (Harm & Seidenberg, 1999; Manis et al., 1996) predicts a smaller than normal regularity effect for phonological dyslexics, but a normal regularity effect for surface dyslexics. Unlike the dual-route model, the connectionist model predicts that the deficits in both subgroups should affect both stimulus types to some extent, because the pronunciations of printed words are learned via a single mechanism. In the case of phonological dyslexics, if the phonological deficit is severe enough, it should affect learning of both exceptions and regular words. In the case of the surface dyslexics, the presence of some type of processing deficit that affects the learning of item-specific word to pronunciation mappings should affect the learning of any new word to some extent (Harm & Seidenberg, 1999). Hence, the connectionist model predicts that both dyslexic subgroups will perform at a lower level than the CA group on both stimulus types. Finally, the connectionist model predicts that surface dyslexics' overall performance on the novel words should be more similar to that of the RL group than is the case for phonological dyslexics.

Method

Participants

The sample of children in the present study was part of a longitudinal study that tested two cohorts totaling 230 children over a 3-year period (see Manis et al., 1999). Teachers were asked to nominate children who were not mentally retarded and who had normal vision and hearing and fluent command of spoken English. Assuming these general inclusion criteria, teachers were then asked to nominate children for the specific groups of interest.

Dyslexics

There were 72 fourth- and fifth-grade dyslexic readers that formed the basis for the subgroups analyzed in the present study. In the first year of the longitudinal study (1996–1997), teachers were asked to nominate children in the third grade who fell within the bottom 25% of their class in reading ability. This was repeated in the second year of the study (1997–1998) so that two cohorts of poor readers, totaling 105 children, were recruited. At the point at which the data for the current study were collected (Fall of 1998), 92 of the poor reader nominees remained in the study, and 72 qualified as dyslexic, based on criteria described subsequently. This dyslexic group consisted of 26 fourth graders and 46 fifth graders. To be classified as dyslexic, children were required to score at or below the 25th percentile on the Woodcock Word Identification Test (Woodcock, 1987), which corresponds to a standard score of 90, and to obtain a standard score of at least 85 (the test sample mean is 100 and the standard deviation is 15) on either the Peabody Picture Vocabulary Test (Dunn & Dunn, 1981) or the Visual Closure subtest of the Woodcock–Johnson Test of Cognitive Ability—revised (Woodcock & Johnson, 1989). A cut-off of 85 on the latter two tests was utilized as a means of ensuring that none of the children had below average intellectual ability.

Reading level-matched comparison group (RL comparison group)

This group consisted of 13 first-, second-, and third-grade students with a mean age of 91.1 months ($SD = 9.5$). Members of the RL comparison group were recruited in 1998 in the spring of their first- or second-grade year, when teachers were asked to nominate average to above average readers. To be included in the RL group, children had to score between the 40th and the 97th percentiles on the Woodcock Word Identification test. In addition to meeting the reading score criteria, participants were held to the same criteria as the dyslexics on the two measures of cognitive ability. The purpose of including this group was to determine whether dyslexics differed from normal readers with a similar overall knowledge of printed words.

Chronological age comparison group

This group consisted of 13 fifth-grade students with a mean age of 125.1 months ($SD = 4.2$) who were recruited 1 year after the end of the longitudinal study (2001) and tested on portions of the same test battery, including the word learning task. These children were nominated by teachers as average to above average readers and had to score between the 40th and the 97th percentile on the Woodcock Word Identification test to be assigned to this group. The criteria for the two measures of cognitive ability were the same as the other groups. All of the referred children met these criteria.

Procedure

In order to identify children that met typical research criteria for surface and phonological dyslexia, participants were assessed on measures of nonsense word reading and exception word reading.

Nonsense word reading task

A list of 70 unfamiliar nonsense words was created for the study. The items were presented in ascending order of difficulty based on pilot data. The items were presented in a notebook in 24-point Arial font in groups of 6 items per page. The items ranged from simple CVC patterns (e.g., *nug*, *bim*) to patterns with two or more letter clusters (e.g., *smaip*, *cleesh*), long vowel patterns (*phuve*) or two syllables (*stining*, *metion*). Children read the items aloud beginning with item 1. Testing was discontinued when children made six consecutive mistakes. Responses were scored correct if the child pronounced all phonemes within the word in accordance with spelling–sound correspondence in English, taking into account the overall orthographic structure of the word. For example, /blet/ was the only possible pronunciation for *blate*. Although *a* can be given a different sound in other orthographic contexts, when followed by a consonant and a final silent *e* it typically is not. In most cases there was a single correct response, but for a few items a second or third response was accepted (e.g., *chome* could be pronounced to rhyme with *dome* or *some*). The items were arranged in ascending order of difficulty based on pilot data. Split-half reliability with a Spearman–Brown correction for length was 0.96.

Exception word reading task

A list of 70 exception words was presented in the same format and font as the nonsense words. There were two types of words on the list. Some words had uncommon spelling patterns that were not pronounced in keeping with typical spelling–sound correspondences (e.g., *people*, *beauty*, *yacht*, *silhouette*). Other words had common spelling patterns with spelling–sound correspondences that ran counter to the typical pronunciation (e.g., *said*, *prove*, *colonel*). The list was ordered from easiest to hardest based on frequency and grade norms, as well as pilot data, and was based in part on a list constructed by Adams and Huggins (1985). Testing was discontinued if the child made six consecutive pronunciation errors. Split-half reliability with a Spearman–Brown correction for length was 0.94.

Phoneme deletion task

A task of the type devised by Bruce (1964) was administered in two parts. In part 1, participants repeated a familiar word that was spoken on a tape. The speaker on the tape asked the participant to repeat the word again, but with a specified part missing, such as "snow" without the /s/ and "act" without the /k/. The items consisted of a single phoneme or a blend of two phonemes that was deleted from the beginning, middle, or end of the word. There were 25 items in part 1, and testing was discontinued if the child made five mistakes in a row. All correct responses were real words, and many of the most common errors were real words. Split-half reliability for word stimuli was 0.88. In part 2, the items were all nonsense words, such as "kimp" without the /m/. There were 15 items, and testing was discontinued if the child made five mistakes in a row. All correct responses were nonsense words. Split-half reliability for nonsense words was 0.87. Scores were combined across the two tasks for the present analyses.

Orthographic choice task

The orthographic choice task (adapted from Olson et al., 1989) required the participant to view two strings of letters displayed just left and just right of a fixation cross on a computer screen and choose the item that was a correctly spelled word. The items were displayed in a large and easily readable font (Arial 24). The participant pressed a button on a box to indicate the side of the screen that contained the correct answer. Half of the items contained an exception word (e.g., *sponge spunge*) and half contained a regular real word (e.g., *sheep sheap*). All of the foil items were phonologically identical to the targeted exemplar item when given their modal pronunciation. Hence the child could not rely on phonological decoding strategies to choose the correct answer. Both paper and computer practice items were administered with corrective feedback to ensure the child understood the task. All 48 test items (and 4 practice items) were administered. For each participant, the order of the stimuli was chosen at random by the computer program. Reaction time was recorded, but not analyzed in the present study owing to fairly low accuracy levels in at least one group. The CA group was not given this task. Split-half reliability was 0.77.

Title recognition task

Participants were shown a list of 45 book titles, including 30 actual book titles and 15 foils, on a sheet of paper, in a task originally created for children by Cunningham and Stanovich (1990). Pilot data from children not in this particular study were collected to ensure that most of these book titles were familiar to at least some children at the school test sites. The titles of books that had been made into movies or are prominent in the school's reading programs were not included in the list. The experimenter read the book titles aloud to the child one at a time while the child looked at a printed list containing the book titles. The child was then asked to decide if each title was a real book or not. Participants were encouraged not to guess and were told some of the book titles were not real. The total score on the task was the proportion of correct book titles minus the proportion of incorrect titles chosen by the child. This score corrects for guessing in the manner devised by Cunningham and Stanovich (1990). The split-half reliability for correct answers, with a Spearman–Brown correction for length, was .80. Two children in the RL group were missing this score.

Novel nonsense word learning task

Participants were trained to pronounce 24 nonsense words. However, 2 items were assigned to the wrong category (that is a regular pronunciation was termed irregular and vice versa (*iche* and *scyfe*)). These items were dropped from analyses. Analyses were conducted only on the remaining 22 items (shown in the Appendix). Participants received a total of six training trials over the 2 training days (three per day). Additionally, participants received two posttest trials over the course of 2 separate testing days. Training and testing days were 3 to 5 days apart and the great majority of children finished all four sessions within a 2-week period.

On the first day of training, participants were shown an outline drawing of a group of "space creatures" taken from the Woodcock–Johnson Cognitive Battery—revised (1989) and instructed that they were going to learn how to read the names of 24 space aliens. They were then shown printed versions of the novel words and trained to pronounce them. None of the novel words were explicitly connected to any of the space creatures, and the picture of the space creatures was not shown again during training. The purpose of the picture was simply to increase the child's interest in the word-learning task. The first trial of each training day was treated as a familiarization trial in which the experimenter showed the child the novel word printed on a sheet of paper, pronounced each printed item aloud, and asked the child to repeat it. Praise was given for correct repetitions of the pronunciation and corrective feedback was given for errors. If the child made a repetition error, the experimenter simply repeated the specific item until the participant could pronounce the item correctly. For the two remaining training trials, children were asked to generate the correct pronunciation on their own from the printed version of the novel word. Feedback and praise were handled in the same manner. There were 2 training days. Thus, accuracy data were recorded for the second and third trials on each training day. Items were simply marked correct or incorrect.

After completing the training, participants had two posttest sessions separated by 3 to 4 days. In the first posttest session (the third session of the 2-week process), participants were given a spelling test. Participants did not receive any additional training on the stimuli. Participants were instructed to spell the novel words on an answer sheet. The experimenter pronounced each item twice for the child and asked the child to spell it on the answer sheet. No feedback was given on this task. The children's spellings were scored later for accuracy and only completely correct spellings were accepted as correct.

In the final session (word-reading posttest), participants attempted to pronounce each novel word again. The novel words were presented on a computer screen. A fixation cross appeared on the screen to signal the onset of each trial. Participants were told that they were going to see the new words on the computer screen, and they should pronounce the items as accurately as they could. Accuracy data were recorded by the experimenter using a button box. Reaction times (from onset of the stimulus to onset of the participant's oral response) were obtained using lapel-mounted microphones connected to the computer through a voice-activated response box. However, the relatively low levels of accuracy on the task and the high within-participant and within-group variability in reaction time limit how informative this type of data can be, so analyses were not carried out.

The stimuli consisted of 22 one-syllable nonsense words ranging from 4 to 6 letters in length (see the Appendix). Each individual nonsense word was assigned one of two possible pronunciations, a regular pronunciation or an exceptional pronunciation. Each participant was taught 11 regular pronunciations and 11 exception word pronunciations over the course of the six training trials. For example, in the regular pronunciation condition, the nonsense word *hape* would be pronounced /hep/, as this pronunciation uses the most frequent pronunciation of the vowel pattern. However, in the exceptional pronunciation condition, it was pronounced /haep/, which runs counter to the most frequent pronunciation (analogous to *have*). There were 12 orthographically regular items like *hape*, which has many orthographic neighbors (*ape*, *cape*, *gape*, etc.). There were 10 "strange" items (e.g., *mouge*, *duite*, *lource*, *choip*) whose exceptional pronunciations have only one or two exemplars in English (e.g., *rouge*, *suite*, *source*, *choice*). For the strange items, the regular pronunciation was based on the most common pronunciation of single-letter or bigram units contained in the words (e.g., /au/ for *ou*). The exceptional pronunciation was based on the pronunciation of the only neighbor of that word in English (e.g., *rouge*), except for one word (*kuise*) that had two neighbors with exceptional pronunciations (*guise* and *disguise*). The novel words were counterbalanced for pronunciation type (regular, exception) across two between-group list conditions so that a word given a regular pronunciation for participants assigned to list 1 had an exception word pronunciation for those assigned to list 2. Six participants from each group were assigned to list 1 and seven to list 2. Orthographically regular and strange items were not perfectly counterbalanced across lists. There were 6 strange items in list 1 and 4 in list 2.

The stimuli appeared in boldface, 64-point type and were centered on laminated sheets of 8.5 by 11-in. white paper organized inside a binder. The items appeared in a different order for each of the six training and two testing trials, and this ordering was constant across participants within a particular list condition. The order of the stimuli for each presentation was quasi-randomly determined to control for pronunciation type (i.e., regular or exception) such that no more than three words of the same type could appear in succession.

Subgroup categorization

Subgroup categorization for the dyslexics was based on a discrepancy score using exception words and nonsense words. We first created a local normative group of 70 fourth and fifth graders out of the sample of 162 children in these grades remaining in the study in Fall, 1998. The normative group was stratified based on percentile scores. The distribution of their percentile scores approximated the distribution of reading scores for national norms on the Woodcock Word Identification test (e.g., about 25% of the normative group scored at or below the 25th percentile and 25% at or above the 75th percentile). Exception word reading and nonsense word reading scores were converted to Z scores based on the normative group's scores. The nonword Z score was subtracted from the exception word Z score to create a discrepancy score. The distribution of the discrepancy scores was approximately normal. On this dimension, low negative scores represented a surface dyslexia profile and high positive scores a phonological dyslexia profile. The top and bottom 25% of this distribution were assigned to subgroups. Because there were 72 dyslexics, the resulting subgroups consisted of 18 children each. The discrepancy score method was used because it encoded differences in nonsense word and exception word reading skill, while allowing for wide variability in reading ability due to individual differences in overall reading impairment within the dyslexic sample (Castles et al., 1999).

After children were categorized as phonological dyslexic or surface dyslexic, they were matched on a case by case basis within 0.5 months on Woodcock Word Identification grade equivalent scores. This process yielded 13 phonological dyslexics with a mean age of 121.4 months ($SD = 5.8$) and 13 surface dyslexics with a mean age of 124.8 months ($SD = 6.99$). The other potential phonological dyslexics and surface dyslexics were not used as they did not match closely enough on Word Identification. Members of the RL comparison group were selected from a pool of 34 children using the same matching procedure, yielding a group of 13 (in order to achieve equal cell size). The RL cases excluded from matching were children who scored either below or above the range of Woodcock Word Identification scores found in the dyslexic samples or who scored more than 2 standard deviations above the dyslexic sample mean on either exception word reading or nonsense word reading. This procedure tended to produce an RL group of average to above average readers, exclusive of extremely gifted readers. The identifying information for the four groups given the word-learning task is given in Table 1.

Results

Initial comparison of the subgroups

To determine whether the groups were appropriately defined, one-way ANOVAs were conducted comparing groups on the reading and cognitive ability measures. Significant ANOVAs were followed by post hoc tests using Tukey HSD, to control

Table 1
Means, standard deviations, and group differences for phonological dyslexics, surface dyslexics, reading level-matched group, and age-matched controls on identifying measures

Variable	Phonological dyslexics ($n = 13$)	Surface dyslexics ($n = 13$)	Reading level comparison group ($n = 13$)	Chronological age-matched comparison group ($n = 13$)
Woodcock Word Identification grade equivalent	3.09^{a} (0.58)	3.02^{a} (0.64)	3.10^{a} (0.58)	6.09^{b} (1.92)
Woodcock Word Identification standard score (mean = 100, $SD = 15$)	76.4^{a} (7.5)	73.9^{a} (9.0)	114.5^{b} (9.7)	102.8^{b} (10.6)
Peabody Picture vocabulary—revised (mean = 100, $SD = 15$)	86.7^{a} (14.6)	91.3^{ab} (17.4)	106.5^{bc} (16.4)	108.8^{c} (14.3)
Woodcock Visual Closure (standard score) (mean = 100, $SD = 15$)	96.1^{a} (12.9)	101.8^{a} (13.1)	105.2^{a} (12.5)	100.9^{a} (14.7)
Title Recognition Task (corrected %)	25.3^{ab} (25.2)	29.5^{ab} (19.6)	14.6^{a} (17.0)	39.9^{b} (12.9)

Mean scores within rows that do not share a common superscript differ significantly in post hoc pairwise comparison analyses, using the Tukey HSD test ($p < .05$).

cumulative type I error. Means, standard deviations, and results of post hoc comparisons are shown in Table 1. Significant group differences were obtained on Woodcock Word Identification grade equivalents, $F(3,48) = 24.92$, $p < .001$, and standard scores, $F(3,48) = 65.96$, $p < .001$, consistent with the definition of the groups. It is notable that phonological dyslexics and surface dyslexics did not differ on grade equivalent or standard scores, and neither group differed from the RL group on the grade equivalent score. Group differences were obtained on Peabody Picture Vocabulary, $F(3,48) = 6.31$, $p < .001$. The phonological dyslexia group differed from the RL group ($p < .025$) and the CA group ($p < .01$) on Peabody Picture Vocabulary, but phonological dyslexics and surface dyslexics did not differ. No group differences on Visual Closure were observed. Thus, the dyslexic subgroups were comparable in reading and cognitive ability.

It is important to establish that the dyslexic subgroups met the criteria in past research studies for phonological and surface dyslexia on key defining and validating tasks. Mean scores for all four groups are presented in Fig. 1. A limited number of planned contrasts using t tests (12 to be exact) with α set at .05 were conducted for the comparisons between the subgroups and the RL group on four key variables: the defining measures, Nonword and Exception Word Reading, and the validating measures, Orthographic Choice and Phoneme Deletion. As expected, phonological dyslexics significantly outperformed surface dyslexics on exception word reading ($p < .025$), whereas this result was reversed on Nonword Reading ($p < .05$). Likewise, phonological dyslexics achieved higher scores on Orthographic Choice ($p < .025$) and surface dyslexics on Phoneme Deletion ($p < .01$), providing some validation for the assignment to subgroups.

Surface dyslexics did not differ significantly from the RL group on Exception Words, Orthographic Choice, Phoneme Deletion, or Nonword Reading, although they showed a trend to score lower than the RL group on Nonword Reading. Surface dyslexics scored lower on all of these tasks than the CA group. The findings are in accord with previous studies. Surface dyslexics appeared to be delayed in word reading generally and appeared to have a mild phonological impairment that could be either a result or a cause of slow reading development (Castles et al., 1999; Manis et al., 1996, 1999; Stanovich et al., 1997). Phonological dyslexics, in contrast, performed less accurately than the RL group on Phoneme Deletion ($p < .01$) and Nonword Reading ($p < .01$), consistent with moderate to severe phonological processing deficits. Although it is apparent that the CA group was superior to the other groups in absolute score, no planned comparisons involving the CA group were conducted.

The similarity of these data to the results of Manis et al. (1996) is striking. Data from the earlier study are shown in Fig. 2 (a figure not included in the original paper). The exception and nonword reading tasks were procedurally similar in the 1996 and current studies, but the exception word items were relatively harder in the 1996 study. The two meaures of phoneme awareness were different (the task was Position Analysis in the earlier study). In both figures, it is apparent that surface dyslexics resemble the RL group closely, whereas the phonological dyslexia group differs in its performance profile from all other groups.

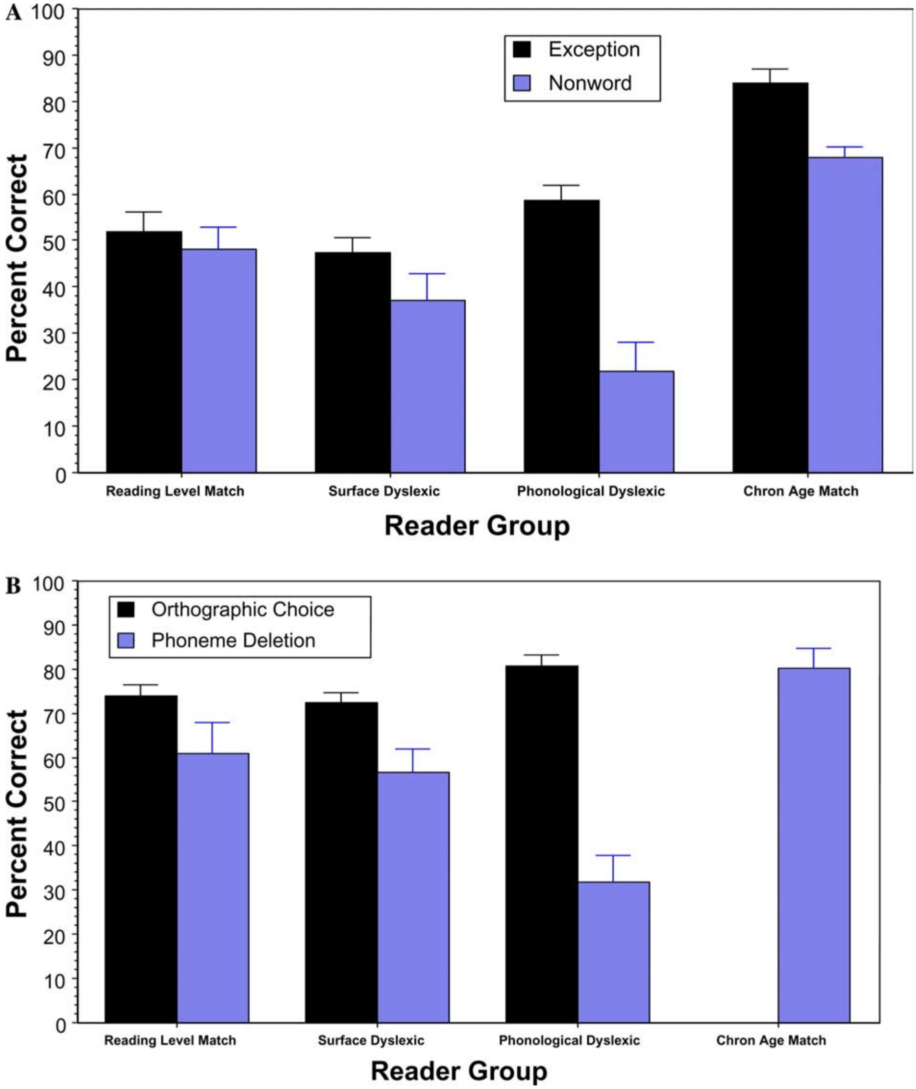

Fig. 1. Mean percentage correct on exception words and nonwords (A) and on Orthographic Choice and Phoneme Deletion (B) for the four groups.

An overall group difference on Title Recognition occurred, $F(3, 46) = 3.13$, $p < .035$ (see Table 1). Phonological dyslexics, surface dyslexics, and RLs did not differ from each other, but the CA group had higher print exposure than the RL group ($p < .025$). The failure to obtain differences between the phonological dyslexia and the surface dyslexia groups on the print exposure measure was contrary to the predictions of Stanovich et al. (1997).

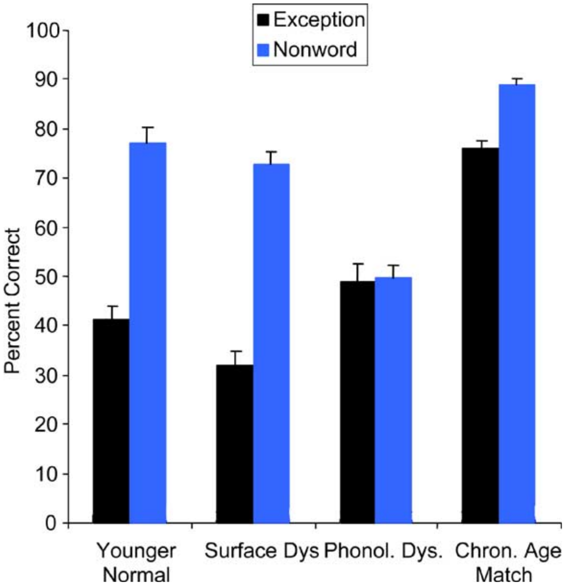

Fig. 2. Mean percentage correct on exception and nonword reading from Manis et al. (1996) (figure not previously published).

Subgroup comparisons on the novel-word-learning task

Preliminary analyses revealed no main effects of stimulus list and no interactions of list with subgroup or regularity. Hence subsequent analyses collapsed across list. The three models make different predictions about the pattern of performance on the novel-word-learning task. The key predictions have to do with the size of the advantage for novel words with regular pronunciations and the difference in performance between normal readers and each dyslexic subgroup. The clearest test of these differential predictions is arguably provided by the word-pronunciation posttest (session 4), because this is a measure of retention of what was learned about the novel words over a period of several days, rather than a short-term effect of training. The data are presented in Fig. 3. The pattern of findings in Fig. 3 resembles the pattern in Fig. 2 fairly closely, even though the procedures and stimuli were different.

It is apparent from the figure that many children succeeded in learning a great number of the novel words. The CA group showed high mastery of both regular and exception novel words. There was a significant main effect for Group, $F(1,48) = 8.28$, $p < .001$ ($\eta^2 = .341$), and for Pronunciation Type, $F(1,48) = 20.09$, $p < .001$ ($\eta^2 = .295$). However, the most interesting finding was an interaction between Group and Pronunciation Type, $F(3,48) = 4.43$, $p < .01$ ($\eta^2 = .217$). The interaction remained significant when the analysis was conducted with only the dyslexic subgroups, $F(1,24) = 9.22$, $p < .01$ ($\eta^2 = .277$). The interaction is clearly interpretable as the absence of a regularity effect in the phonological dyslexia group alone. To confirm this finding, simple main effects of regularity were analyzed for each group. The

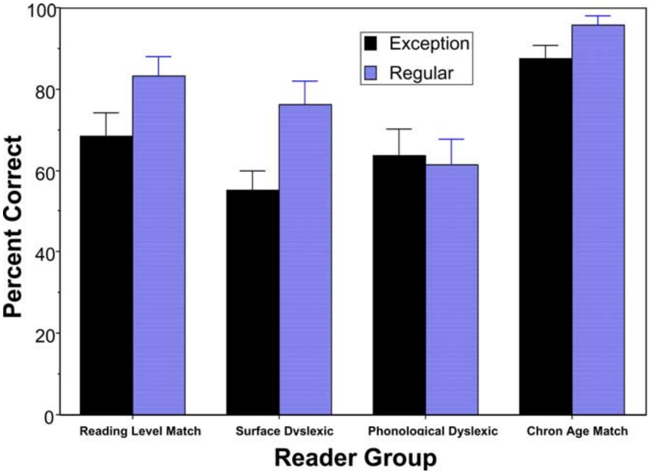

Fig. 3. Mean percentage correct for regular and exception items on the novel-word-pronunciation posttest for all four groups.

Pronunciation Type effect was significant for surface dyslexics, $F(1,12) = 17.03$, $p < .001$; for the CA group, $F(1,12) = 12.07$, $p < .01$; and for the RL group, $F(1,12) = 9.031, p < .025$; but not for the phonological dyslexia group. The significant interaction runs counter to the hypothesis of Stanovich et al. (1997), which views both dyslexic subgroups as having a phonological deficit.

Planned comparisons using ANOVA were conducted between the phonological dyslexia and the surface dyslexia groups and between the dyslexic subgroups and each of the normal reader groups, for regular novel words and for exceptional novel words. There were eight tests conducted, with a p value set at .05 for each test. Phonological dyslexics and surface dyslexics did not differ significantly on either regular or exceptional novel words, although the direction of differences favored surface dyslexics on the regular and phonological dyslexics on the exceptional novel words. The phonological dyslexia group performed less accurately than the CA group on both exceptional novel words, $F(1,24) = 10.26$, $p < .01$, and regular novel words, $F(1,24) = 26.7$, $p < .001$. Differences between the phonological dyslexia and the RL groups were much smaller, and only the difference on regular novel words was reliable, $F(1,24) = 7.65, p < .025$. The surface dyslexia group also performed less accurately than the CA group on both exceptional novel words, $F(1,24) = 30.82$, $p < .001$, and regular novel words, $F(1,24) = 10.36, p < .01$. No differences between the surface dyslexia and the RL groups were obtained. This pattern of differences between the dyslexics and the normal reader groups is inconsistent with the dual-route model, which predicted normal levels of performance for the phonological dyslexia group on novel words with exceptional pronunciations and for the surface dyslexia group on novel words with regular pronunciations. Differences between the phonological dyslexia group and the RL group on the regular novel words were consistent with a severe phonological processing problem. The pattern of results in the word-learning task bears a strong resemblance to the data from traditional exception and nonword measures (see Figs. 1 and 2), indicating developmental deviance in the phonological dyslexia group and developmental delay in the surface dyslexia group.

It is possible that some of the group differences in printed word learning were attributable to differences in verbal ability. Peabody Picture Vocabulary Test—revised (PPVT-R) scores were correlated with the novel-word-pronunciation posttest (.46 with regular word pronunciation and .40 with exception word pronunciation for the combined sample of 52 participants). An ANCOVA was conducted on the novel-word-pronunciation posttest with PPVT-R as a covariate. PPVT-R did not account for reliable between-participants variance in this analysis ($\eta^2 = .075$), and the Group main effect was still significant, $F(3, 47) = 4.36$, $p < .01$ ($\eta^2 = .218$). In addition, PPVT-R did not interact with Pronunciation Type, making it unlikely that the Group by Pronunciation Type interaction was an artifact of differences in overall verbal ability.

Training trials

It is of interest whether the word reading profiles on the posttest were present earlier in training. None of the theories necessarily predicted a Group by Pronunciation Type by Trial interaction, particularly over such a short period of learning (six training trials). However, there is some basis for expecting an interaction of Group by Trial. That is, dyslexics might require more feedback to reach the same level of performance as the control groups.

The first two sessions were training sessions, with three trials per session. Because the first trial in each session did not involve a participant-generated pronunciation, there were two trials per session that could serve as data on learning rate (a total of four learning trials). The data for the second and third trials of session 1 and the second and third trials of session 2, which are referred to here for convenience as trials 1–4, are shown in Fig. 4. An ANOVA with factors of Group (four levels),

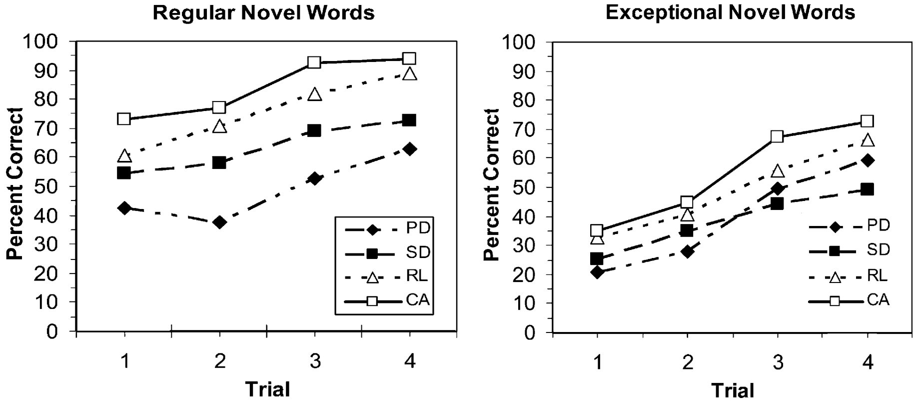

Fig. 4. Mean percentage correct as a function of training trial for the four groups: (left) novel words with regular pronunciations and (right) novel words with exceptional pronunciations.

Pronunciation Type, and Trial revealed a significant Group effect, $F(3,48) = 12.66$, $p < .001$ ($\eta^2 = .442$), with both normal reader groups scoring at a higher level across trials than the dyslexic subgroups. There were also main effects of Pronunciation Type, $F(1,48) = 128.5$, $p < .001$ ($\eta^2 = .728$), due to the superiority of novel words with regular pronunciations, and Trial, $F(3,46) = 117.05$, $p < .001$ ($\eta^2 = .709$), due to increasing accuracy across trials. The linear component of Trial was significant, $F(1,48) = 236.3$, $p < .001$ ($\eta^2 = .831$), as was the cubic, $F(1,48) = 13.38$, $p < .001$ ($\eta^2 = .218$). Though clearly the linear trend is the stronger of the two effects, the cubic trend reflects some slowing of growth (or ceiling effects) on the later trials. Group interacted with Pronunciation Type, $F(3,48) = 5.00, p < .01$ ($\eta^2 = .238$), and Pronunciation Type interacted with Trial, $F(3,144) = 4.43$, $p < .01$ ($\eta^2 = .085$). There were no other significant interactions. In an analysis repeated with only phonological dyslexics and surface dyslexics, the Group effect was not significant but the Group by Pronunciation Type interaction remained significant, $F(1,24) = 7.34$, $p < .025$ ($\eta^2 = .234$), indicating that subgroup differences were a major source of this interaction. The Pronunciation Type by Trial interaction also remained significant in the analysis with only phonological dyslexics and surface dyslexics, $F(3,72) = 3.01$, $p < .05$ ($\eta^2 = .112$).

Fig. 4 helps explicate the interactions. All four groups showed better performance on novel words with regular pronunciations than on novel words with exceptional pronunciations, accounting for the Pronunciation Type effect, but this difference was sustained across trials for surface dyslexics, RLs, and CAs, whereas it was present only on trial 1 for the phonological dyslexics. This is consistent with the observed Group by Pronunciation Type interaction. The Pronunciation Type by Trial interaction is readily apparent in Fig. 4 as a steeper rate of growth for the exceptional novel words than for the regular novel words, collapsing across groups. This may be due to the lower baseline level of performance on the exceptional items (i.e., there was more room for improvement). The absence of a Group by Trial interaction indicates that the groups learned the novel words at a similar rate, even though dyslexics obviously attained lower overall levels of accuracy. The absence of a Group by Pronunciation Type by Trial interaction, and a careful inspection of the data, indicated that neither of the dyslexic subgroups showed dramatically slower learning of exceptional novel words than regular novel words.

To evaluate the contribution of PPVT-R scores to word learning across trials, we conducted an ANCOVA on the training trials with PPVT-R standard score as a covariate. The Group main effect was still significant, $F(3,47) = 7.53$, $p < .001$ ($\eta^2 = .325$), and Peabody scores did not interact with Pronunciation Type or Trial.

Spelling posttest

Fig. 5 depicts the spelling data (from session 3). An ANOVA with factors of Group and Pronunciation Type revealed a significant group effect, $F(3,48) = 28.66, p < .001$ ($\eta^2 = .642$), which appeared to be due largely to higher performance by the CA group. In fact, the other three groups spelled less than 40% of the regular novel words, and less than 26% of the exceptional novel words, correctly. The interaction between Group and Pronunciation Type was also significant, $F(3,48) = 2.84$, $p < .05$ ($\eta^2 = .151$).

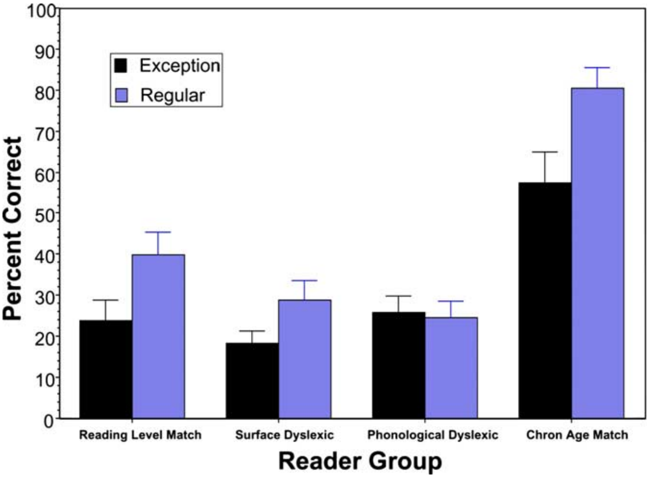

Fig. 5. Mean percentage correct for novel words with regular and exception pronunciations on the spelling posttest for all four groups.

Tests of simple main effects of Pronunciation Type for each group indicated a significant regularity effect for the RL group, $F(1,12) = 9.71$, $p < .01$, and the CA group, $F(1,12) = 7.99$, $p < .025$, but not for the two dyslexic subgroups. The Group effect was still significant in an ANCOVA when PPVT-R was entered as a covariate, $F(3,47) = 23.52$, $p < .001$ ($\eta^2 = .600$), and PPVT-R did not interact with Pronunciation Type.

A limited number of planned comparisons were conducted using ANOVAs between the phonological dyslexia and the surface dyslexia groups and between the dyslexic subgroups and each of the normal reader groups. Phonological dyslexics and surface dyslexics did not differ significantly on either regular or exceptional novel words, although the direction of differences favored surface dyslexics on the regular and phonological dyslexics on the exceptional novel words. The phonological dyslexia group performed more poorly than the CA group on both regular novel words, $F(1,24) = 73.02$, $p < .001$, and exceptional novel words, $F(1,24) = 13.44$, $p < .01$. Differences between the phonological dyslexia and the RL groups were reliable only on regular items, $F(1,24) = 5.09$, $p < .025$. The surface dyslexia group also performed more poorly than the CA on both regular items, $F(1,24) = 55.06$, $p < .001$, and exception items, $F(1,24) = 22.76$, $p < .001$. No differences between the surface dyslexia and the RL groups were obtained. The spelling data showed some of the same trends as the word-pronunciation posttest, but the extreme difficulty of the task may have made it less sensitive to group and pronunciation type differences.

One objection that can be raised to the study design is that the phonological dyslexia and surface dyslexia groups were not composed of "pure" cases, and hence the study is not a true test of the dual-route framework. Such extreme profiles are rare among developmental dyslexics. There were five cases in each of the phonological dyslexia and surface dyslexia subgroups that met the criteria for pure subgroups used in past studies (Castles & Holmes, 1996; Murphy & Pollatsek, 1994). All five cases in each group were within 1 standard deviation of the local norm group for their age level (third–fourth graders) on one of the defining measures (nonword or exception word reading) but below the normal range on the other measure. Fig. 6 shows the data for the pure cases on the word pronunciation and spelling posttests. It can be seen that the "pure" subgroups performed slightly better than the full subgroups, but the data correspond closely to the overall pattern present in Figs. 3 and 5 for the full subgroups.

Discussion

The results provide evidence that surface and phonological dyslexics are different, using a novel-word-learning methodology, rather than the overt word and nonword naming tasks used in most previous studies. We focus here primarily on the word-pronunciation data (Figs. 3 and 4) rather than the spelling data, because of possible floor effects in the latter measure. In addition to differences between the subgroups, the fact that phonological dyslexics differed from both the CA and the RL groups, and surface dyslexics differed only from the CA group (see Figs. 1, 3, and 4), leads us to infer that phonological dyslexics develop along a different developmental path than normal readers of the same overall reading level. In contrast, surface dyslexics

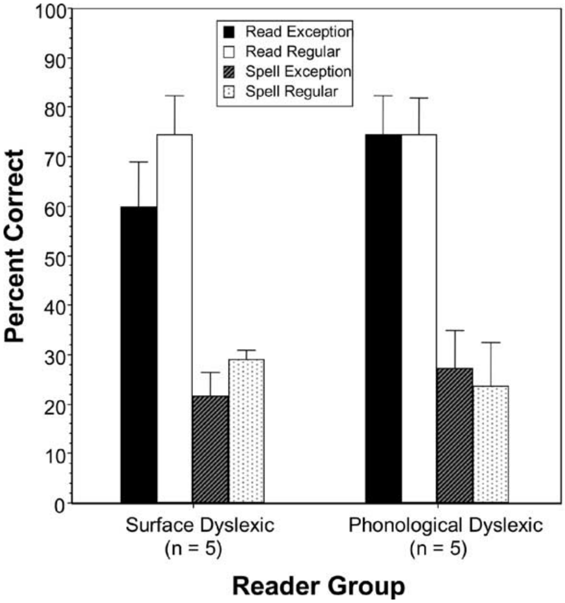

Fig. 6. Mean percentage correct for regular and exception items on the word-pronunciation posttest for two "pure" subgroups.

resemble younger normal readers closely (see Figs. 1, 3, and 4), and hence appear to have a type of deficit in reading that causes a general retardation in the development of knowledge and skill in dealing with printed words, rather than specific deficits in phonological processing.

The study replicated and extended a printed-word-learning study by Castles and Holmes (1996). In addition, the pattern of findings is in close accord with results obtained from word- and nonword-pronunciation tasks (Manis et al., 1996; Stanovich et al., 1997) (compare Figs. 1–3). This justifies a general conclusion, which is that researchers should be looking carefully at differences among dyslexics, rather than treating them as a single group. We evaluated three explanations for the phonological and surface dyslexic patterns.

Stanovich et al. (1997) proposed that the subgroups could be explained on the basis of two factors, degree of phonological impairment and amount of print exposure. This view predicts that both dyslexic subgroups would be somewhat insensitive to spelling-sound regularity in the trained items. This was not the case. Only the phonological dyslexia group was insensitive to regularity. The surface dyslexia group was as sensitive as the normal reader groups (see Fig. 3). In addition, the argument that surface dyslexics have low print exposure was not borne out by the title recognition task that was administered. Hence the Stanovich et al. (1997) hypothesis does not account well for the present data. In addition, the existence of a small number of dyslexics reported in previous group and case studies who were completely normal in phonological skills (e.g., Castles & Coltheart, 1993, 1996; Manis et al., 1996, 1999; Stanovich et al., 1997) (represented here by the five "pure" surface dyslexia cases in Fig. 6) is not explained well by this view. However, the observation that surface dyslexics showed a nonsignificant trend to perform more poorly than the RL group on the nonsense word and phoneme deletion tasks (see Fig. 1) is consistent with the argument by Stanovich et al. (1997) that surface dyslexics have a mild phonological deficit. The problem with such a finding in cross-sectional data such as the present set of data is that it could be either a result or a cause of poor reading. An analysis of the full sample of participants involved in the present study over a 2-year period by Manis et al. (1999) revealed that surface dyslexics had a phoneme awareness deficit relative to the RL group that ameliorated by grade 4 (the year before the present data collection for fifth graders in the present study and the year of the present data collection for fourth graders in the present study). This suggests that mild phonological deficits may indeed be a part of the developmental profile of surface dyslexics, perhaps earlier in development. The important point for present purposes is that mild phonological deficits alone cannot account for the profile of poor exception word reading relative to nonword and regular word reading and for the fact that surface dyslexics and phonological dyslexics were similar in Woodcock word identification skill.

The dual-route model (Castles & Coltheart, 1993; Coltheart et al., 1993) proposes that the sublexical and lexical reading mechanisms develop independently and are impaired independently. This model does not account easily for the substantial number of dyslexics in previous studies that show impairments in both nonword and exception word reading. In addition, the model predicts that phonological dyslexics would be less accurate than the CA group at learning to read novel regular words, but not novel exception words. In fact, the phonological dyslexics were less accurate than the CA group on both pronunciation types. Finally, the model predicts that surface dyslexics should be less accurate than the CA group on exceptional novel words, but not regular novel words, and in fact the surface dyslexics were less accurate on both (see Figs. 3 and 5). The model does not make specific predictions contrasting dyslexic subgroups with the RL group.

The connectionist model (Harm & Seidenberg, 1999; Manis et al., 1996) explains aspects of the data that are problematic for the other approaches. It predicts an interaction between group and word type, with phonological dyslexics showing a smaller regularity effect than surface dyslexics, and this was observed (Fig. 3). In addition, the presence of a phonological impairment should reduce the ability to learn both word types, a prediction that is consistent with the observed difference between the phonological dyslexia and the CA group. The surface dyslexics should also have difficulty coming up to the normal level on both regular and exception words, due to their difficulties with item-specific mappings, and this in fact was observed. In addition, unlike the Stanovich et al. (1997) view, this model does not necessarily predict overall differences between subgroups in print exposure.

An important distinction between the dual-route and the connectionist models is that the connectionist model specifies learning mechanisms explicitly, whereas the mechanism by which children learn to utilize the lexical and nonlexical mechanisms in the dual-route model is not clear. It is possible for dual-route theorists to characterize the nonlexical mechanism as rule learning and the lexical as rote learning (Coltheart et al., 1993), but it is not clear how a child would know which type of mechanism to use with printed words, because words are not labeled in advance as to whether they are regular words, exceptions, or nonsense words. A second difference is that the dual-route model has difficulty explaining the fact that for both children and adults, the ordering of item difficulty across several studies, including the present one, is regular < irregular < nonwords. Regular words should be equal to nonwords if they are read using the nonlexical mechanism or, instead, to exception words if they are read using the lexical mechanism. In the connectionist models, regular words are easier than irregular words because the developing system can draw on previous activations of hidden units and connections for similarly spelled and pronounced words in the case of regular words. Nonwords are hardest of all because the system must make generalizations from existing connections. It is still more difficult to explain the delayed reading profile of the surface dyslexics using the dual-route model, whereas the Harm and Seidenberg (1999) connectionist model provides a parsimonious explanation for both the phonological dyslexia and the surface dyslexia patterns at the level of word reading.

One of the findings of the earlier study by Castles and Holmes (1996) raises a point that has not been previously considered in this paper. Castles and Holmes (1996) reported subgroup differences in word pronunciation accuracy for strange words (e.g., *macht* pronounced as rhyming with *yacht*), but not for what they called exception words. The latter were nonsense words that had common spelling patterns, but were given less common pronunciations, such as *chove*, pronounced by analogy to *dov*e, or *bouch*, pronounced by analogy to *touch*. This raises the possibility that we would find more extreme deficits for surface dyslexics on strange words. Because of the deletion of two stimuli, only 10 of our 22 words were strange words (e.g., *duite, tauge, lource*) and the assignment of words to lists was not perfectly counterbalanced. One list had 4 and the other had 6 strange words with exceptional pronunciations. Nevertheless, collapsing across list, the mean scores reported in Table 2 reveal that the exception/strange distinction does not produce a pattern for either phonological dyslexics or surface dyslexics different from that seen when the two word types are pooled.

It is difficult to characterize the nature of the deficit that gives rise to surface dyslexia more precisely at this point, other than to propose that it is something that affects the encoding of individual exemplars of printed words. Harm and Seidenberg (1999) produced the surface dyslexic pattern in their simulation study either by reducing hidden units or by altering the rate of learning orthographic to phonological correspondences. It is not clear what the behavioral equivalent of reduced hidden units might be, but its effect is to make item-specific learning more difficult. As far as evidence for slower learning rates, the present study found no difference in the rate of learning to associate printed words with their pronunciations in the surface dyslexia group, but the training period was relatively short. Still another explanation might revolve around deficits in the orthographic units. Seidenberg and McClelland (1989) simulated surface dyslexia with damage to the orthographic units, but neither this model nor subsequent models contain mechanisms for learning orthographic regularities. There are data consistent with deficits among surface dyslexics in some aspect of visual or orthographic processing of printed words (Goulandris & Snowling, 1991; Manis et al., 1999). For example, in the latter study, surface dyslexics were slower than phonological dyslexics to decide whether a 5-letter array contained 2 letters of the same name (e.g., gNbGc) or not (HtMdv). At a different level of analysis, some investigators are exploring the possibility that deficits in the processing of orthography are correlated with visual magnocellular channel functioning (Cornelissen & Hansen, 1998; Sperling, Lu, Manis, & Seidenberg, 2003; Talcott et al., 2001) although there are negative findings (Hayduk, Bruck, & Cavanagh, 1996; Spinelli et al., 1997; Williams, Stuart, Castles, & McAnally, 2003).

The failure to find a regularity effect in the phonological dyslexia group is worth noting, given the observation that most groups of dyslexics show regularity effects, as

Table 2
Means and standard deviations for percentage correct on exception novel words (e.g., *zide* → /zid/) and strange novel words (e.g., *duite* → /dut/) on the word-pronunciation posttest for the dyslexic subgroups

Variable	Phonological dyslexics ($n = 13$)	Surface dyslexics ($n = 13$)
Exception novel words	62.4 (24.0)	53.7 (20.2)
Strange novel words	64.8 (32.4)	57.0 (27.6)
Exception plus strange novel words (see Fig. 3)	63.5 (24.1)	55.2 (17.2)
All regular novel words (see Fig. 3)	61.5 (22.6)	76.2 (20.5)

indicated by a review of the literature conducted by Metsala, Stanovich, and Brown (1998). The implication of our data is that some dyslexics are so deficient in phonological decoding that they show no advantage in reading regular words, at least not in their initial attempts to learn particular printed words. This highlights the need to study individual differences in word processing mechanisms among dyslexics, rather than treating them as an undifferentiated group.

In conclusion, our results provide converging evidence from a word-learning task of different developmental profiles among dyslexics. Thus, the profiles are robust across several tasks (word pronunciation, word learning, and orthographic choice, to mention three prominent tasks) (Castles et al., 1999; Manis et al., 1996, 1999; Stanovich et al., 1997). The data differentiated between three competing accounts of these differences between dyslexics. The connectionist model accounted for the results quite well. It provides plausible and testable mechanisms by which printed words might be learned by the reader. In addition, it handles findings from previous studies, such as the prevalence of the mixed dyslexic profile. Hence, we think this type of model, perhaps with modifications to the orthographic units, and implementation of both orthographic to phonological and orthographic to semantic systems (e.g., Harm & Seidenberg, in press) holds a great deal of promise for improving our understanding of reading in general and dyslexia in particular.

Acknowledgments

This research was supported by Grant NICHD 29891 to the second author and Grants NIMH P01-47566 and MH K02-0-1188 to the fourth author. We thank the following people for assistance in collecting and scoring data: Leslie Cunningham, Suzanne Curtin, Melissa Fuess, Mandy Guardiani, Janelle Healy, Megumi Ichikawa, Marc Joanisse, Janeanne Martino, Anne Marie Medina, and Kelsey Willits. We are grateful to the Long Beach Unified School District and to the many teachers, principals, staff, and students who participated in the study.

Appendix

Novel word items and pronunciation guide

Item	Regular pronunciation (and rhyme)	Exception pronunciation (and rhyme)
1. *trom*	/tram/ “prom”	/trom/ “home”
2. *zide*	/zayd/ “ride”	/zId/ “lid”
3. *veep*	/vip/ “jeep”	/vIp/ “zip”
4. *duite*	/dut/ “suit”	/dwit/ “suite”
5. *tauge*	/tadj/ “dodge”	/tedj/ “gauge”
6. *lource*	/laurs/ “sour” + /s/	/lors/ “source”
7. *stoom*	/stum/ “broom”	/stəm/ “some”

Appendix A (*continued*)

Item	Regular pronunciation (and rhyme)	Exception pronunciation (and rhyme)
8. *brote*	/brot/ “wrote”	/brət/ “but”
9. *grast*	/graest/ “fast”	/grest/ “waste”
10. *mouge*	/maudj/ “gouge”	/mudj/ “rouge”
11. *torps*	/torps/ “corpse”	/tor/ “corps”
12. *kuise*	/kuz/ “bruise”	/gayz/ “guise”
13. *dieve*	/div/ “believe”	/dIv/ “sieve”
14. *sloam*	/slom/ “home”	/slam/ “glom”
15. *choip*	/čoip/ “boy” + p	/kwayp/ “swipe”
16. *plish*	/plIš/ “dish”	/playš/ “sly” + /š/
17. *mearse*	/mirs/ “pierce”	/hərs/ “hearse”
18. *smune*	/smun/ “moon”	/smən/ “sun”
19. *froupe*	/fraup/ “brow” + /p/	/frup/ “coupe”
20. *cheam*	/čim/ “beam”	/čaem/ “jam”
21. *hape*	/hep/ “tape”	/haep/ “tap”
22. *glait*	/glet/ “slate”	/glEt/ “wet”

References

Adams, M. J., & Huggins, A. W. F. (1985). The growth of children’s sight word vocabulary: A quick test with educational and theoretical implications. *Reading Research Quarterly, 20*, 262–281.

Backman, J. E., Mamen, M. E., & Ferguson, H. B. (1984). Reading level design: Conceptual and methodological issues in reading research. *Psychological Bulletin, 96*, 560–568.

Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading–spelling patterns. *Developmental Medicine and Child Neurology, 15*, 663–687.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*, 419–421.

Bruce, L. J. (1964). The analysis of word sounds by young children. *British Journal of Educational Psychology, 34*, 158–170.

Bryant, P. E., & Impey, L. (1986). The similarity between normal readers and developmental and acquired dyslexics. *Cognition, 24*, 121–137.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149–180.

Castles, A., & Coltheart, M. (1996). Cognitive correlates of developmental surface dyslexia: A single case study. *Cognitive Neuropsychology, 13*, 25–50.

Castles, A., Datta, H., Gayan, J., & Olson, R. K. (1999). Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology, 72*, 73–94.

Castles, A., & Holmes, V. M. (1996). Subtypes of developmental dyslexia and lexical acquisition. *Australian Journal of Psychology, 48*, 130–135.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). London: Academic Press.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual route and parallel distributed processing approaches. *Psychological Review, 100*, 589–608.

Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *Quarterly Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 718–729.

Cornelissen, P. L., Hansen, P. C., et al. (1998). Motion detection, letter position encoding and single word reading (1998). *Annals of Dyslexia, 48*, 29–62.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology, 28*, 733–740.

Doehring, D. G., Trites, R. I., Patel, P. G., & Fiedorowicz, C. A. M. (1981). *Reading disabilities: The interaction of reading, language, and neuropsychological deficits*. New York: Academic Press.

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

Ehri, L. C., Wilce, L. S., & Taylor, B. B. (1987). Children's categorization of short vowels in words and the influence of spellings. *Merrill–Palmer Quarterly, 33*, 393–421.

Frith, U. (1985). The usefulness of the concept of unexpected reading failure: Comments on "reading retardation revisited". *British Journal of Developmental Psychology, 3*, 15–17.

Goulandris, N. K., & Snowling, M. (1991). Visual memory deficits: A plausible cause of developmental dyslexia. Evidence from a single case study. *Cognitive Neuropsychology, 8*, 127–154.

Hanley, R., Hastie, K., & Kay, J. (1992). Developmental surface dyslexia and dysgraphia: An orthographic processing impairment. *Quarterly Journal of Experimental Psychology, 44*, 285–319.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading and dyslexia: Insights from connectionist models. *Psychological Review, 106*, 491–528.

Harm, M. W., & Seidenberg, M. S. (in press). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*.

Hayduk, S., Bruck, M., & Cavanagh, P. (1996). Low-level visual processing skills of adults and children with dyslexia. *Cognitive Neuropsychology, 13*, 975–1015.

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia, 45*, 3–27.

Lyon, G. R., & Watson, B. (1981). Empirically derived subgroups of learning disabled readers: Diagnostic characteristics. *Journal of Learning Disabilities, 14*, 256–261.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the basis of two subtypes of developmental dyslexia. *Cognition, 58*, 157–195.

Manis, F. R., Seidenberg, M. S., Stallings, L., Joanisse, M. F., Bailey, C. E., Freedman, L. B., & Curtin, S. (1999). Development of dyslexic subgroups: A one year follow up. *Annals of Dyslexia, 49*, 105–136.

Metsala, J. L., Stanovich, K. E., & Brown, G. D. A. (1998). Regularity effects and the phonological deficit model of reading disabilities: A meta-analytic review. *Journal of Educational Psychology, 90*, 279–293.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323–331.

Morris, R., Blashfield, R., & Satz, P. (1981). Neuropsychology and cluster analysis: Problems and pitfalls. *Journal of Clinical Neuropsychology, 3*, 79–99.

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., Katz, L., Francis, D. J., & Shaywitz, B. A. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology, 90*, 347–373.

Murphy, L., & Pollatsek, A. (1994). Developmental dyslexia: Heterogeneity without discrete subgroups. *Annals of Dyslexia, 44*, 120–146.

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

Patterson, K. E., & Morton, J. (1985). From orthography to phonology: An attempt at an old interpretation. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia*. London: Erlbaum.

Pennington, B. (1999). Toward an integrated understanding of dyslexia: Genetic, neurological, and cognitive mechanisms. *Development and Psychopathology, 11*, 629–654.

Perfetti, C. A., Beck, I. L., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill–Palmer Quarterly, 33*, 283–319.

Plaut, D., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Visual word recognition: Are two routes really necessary? *Psychological Review, 103*, 56–115.

Seidenberg, M. S. (1993). A connectionist modeling approach to word recognition and dyslexia. *Psychological Science, 4*, 299–304.

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

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23*, 383–404.

Shankweiler, D., Crain, S., Katz, L., Fowler, A. E., Lieberman, A. E., Brady, S. A., Thornton, R., Lundquist, E., Dreyer, L., Fletcher, J. M., Stuebing, K. K., Shaywitz, S. E., & Shaywitz, B. A. (1995). Cognitive profiles of reading disabled children: Comparison of language skills in phonology, morphology and syntax. *Psychological Science, 6*, 149–156.

Share, D. L. (1995). Phonological recoding and self-teaching: The sine 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 in a model of acquisition. *Issues in Education: Contributions from Educational Psychology, 1*, 1–57.

Sperling, A. J., Lu, A. L., Manis, F. R., & Seidenberg, M. S. (2003). Selective magnocellular deficits in dyslexia: A "phantom contour" study. *Neuropsychologia, 41*, 1422–1429.

Spinelli, D., Angelelli, P., De Luca, M., Di Pace, E., Judica, A., & Zoccolotti, P. (1997). Developmental surface dyslexia is not associated with deficits in the transient visual system. *NeuroReport, 8*, 1807–1812.

Sprenger-Charolles, L., Cole, P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia: Evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology, 54*, 87–104.

Stanovich, K. E. (1988). Explaining differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21*, 590–612.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89*, 114–128.

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

Talcott, J. B., Witton, C., McLean, M. F., Hansen, P. C., Rees, A., Green, G. G. R., & Stein, J. F. (2001). Dynamic sensory sensitivity and children's word decoding skills. *Proceedings of the National Academy of Science, 97*, 2952–2957.

Temple, C. M., & Marshall, J. C. (1983). A case study of developmental phonological dyslexia. *British Journal of Psychology, 74*, 517–533.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Williams, M. J., Stuart, G. W., Castles, A., & McAnally, K. I. (2003). Contrast sensitivity in subgroups of developmental dyslexia. *Vision Research, 43*, 467–477.

Woodcock, R. W. (1987). *Woodcock Reading Mastery Test—revised.* Circle Pines, MN: American Guidance Service.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson Tests of Cognitive Ability—revised.* Allen, TX: DLM Teaching Resources.