Available online at www.sciencedirect.com

ScienceDirect

Journal of Experimental Child Psychology 95 (2006) 153–193

Journal of Experimental Child Psychology

www.elsevier.com/locate/jecp

Deficient implicit phonological representations in children with dyslexia

Richard Boada [*], Bruce F. Pennington

Department of Psychology, University of Denver, Denver, CO 80208, USA

Received 10 March 2004; revised 25 March 2006
Available online 2 August 2006

Abstract

This study tested the segmentation hypothesis of dyslexia by measuring implicit phonological representations in reading-disabled 11- to 13-year-olds. Implicit measures included lexical gating, priming, and syllable similarity tasks designed to reduce metalinguistic demands. Children with dyslexia performed consistently worse than CA and RA controls when more segmental representations were required across all three tasks. Implicit phonological representations were correlated with measures of speech perception, phoneme awareness, and phonological short-term memory, but not rapid automatized naming, and accounted for unique variance in predicting reading ability. Results provide\ strong support for less mature implicit phonological representations in children with dyslexia.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Dyslexia; Implicit phonological representations; Segmentation; Speech perception; Phoneme awareness; Auditory processing; Literacy

Introduction

Overview

It is becoming increasingly clear that early phonological skill in general, and phoneme awareness (PA) skill more specifically, predicts later reading ability (Bird, Bishop, & Freeman, 1995; Olson, Kliegl, Davidson, & Foltz, 1985; Olson, Wise, Conners, Rack, & Fulker, 1989; Pennington & Lefly, 2001; Scarborough, 1990, 1998; Stanovich, 1988; Wagner

---

[*] Corresponding author. Fax: +1 303 871 3982.
E-mail address: rboada@nova.psy.du.edu (R. Boada).

doi:10.1016/j.jecp.2006.04.003

& Torgesen, 1987). Although various hypotheses have been put forth to explain the PA deficit in children with dyslexia, the segmentation hypothesis (Bird & Bishop, 1992; Fowler, 1991) holds promise. This hypothesis requires a deficit in implicit phonological representations, but such a deficit has been difficult to measure. Traditionally, speech perception tasks have been used in an attempt to measure implicit phonological representations; however, the results of such experiments have not been conclusive. This study was undertaken to provide converging evidence, using three experimental tasks with different surface characteristics, of a deficit in implicit phonological representations in children with reading disability.

As stated earlier, researchers have clearly shown that there is a deficit in PA in children and adults with dyslexia, even when compared with reading age-matched participants. Because the majority of these participants did not have clear evidence of gross linguistic deficits, a metalinguistic hypothesis emerged, stating that the development of underlying phonological representations is normal in these individuals, but access to them in an explicit manner for the purposes of learning how to read and spell is problematic. Researchers from different fields subsequently began to document various early articulation, syntactic, and morphological difficulties in some children ultimately diagnosed with dyslexia (e.g., Lewis & Freebairn, 1992; Scarborough, 1990), potentially implicating a deficit below that of PA for at least a subset of children with this disorder. This has led to a concerted effort to identify what lies beneath the level of the PA deficit. Two main theories have been proposed. The first, based on early work by Tallal (1980), proposes a nonspeech auditory perceptual deficit that undermines speech perception and the development of underlying phonological representations and ultimately the development of PA. A second theory, often referred to as the segmentation hypothesis (Bird & Bishop, 1992; Fowler, 1991), proposes that abnormal development of the underlying phonological representations themselves is the primary cause of the subsequent PA deficit; it does not invoke the presence of a more general auditory perceptual deficit. The current study concerns itself primarily with the segmentation hypothesis by attempting to show that there are measurable deficits in phonological representations below the level of PA in children with dyslexia.

The segmentation hypothesis is best understood in the context of normal phonological development. When learning a language, a child builds cognitive–linguistic representations of varying grain sizes, including phrases, words, syllables, demi-syllables, and eventually phonemes (Fowler, 1991; Walley, 1993). This is particularly difficult because there are no invariant acoustic or temporal cues in the speech stream that mark these units, particularly at the sublexical level. The young listener must learn to discern, weigh, and then integrate various acoustic properties along the temporal and spectral domains to derive lexical- and phonological-level representations. This task is mastered slowly during the course of development, leading to measurable differences in speech perception and production tasks between adults and children of various ages (Nittrouer, 1992; Nittrouer & Studdert-Kennedy, 1987). Delays or deviations in this developmental process may lead to, or be associated with, problems with other demanding linguistic tasks such as expanding vocabulary, increasing short-term memory capacity, and becoming aware of phonemes.

Studying the emergence of phonological or cognitive representations is challenging for a number of reasons. There are general-level confounds, such as attention and conceptual ability, that may be different in selected populations. It is also difficult to ensure that one is measuring a discrete level of cognitive–linguistic processing because there always is the possibility of dynamic interaction between top-down and bottom-up processes on performance. More fundamentally, however, cognitive scientists have been unable to specify the exact nature of a phonological representation, even though the term is used often enough in theoretical models, making it difficult to operationalize specific measures of such a construct. At a general level, we know that with appropriate experience with a native language, speech input and gestural (articulatory) output in a child coordinate in such a way as to allow for the emergent capability to represent and ultimately access phonemic structure. The exact process or the precise cues that are used to make this happen, however, are not yet well specified. Thus, we acknowledge that the term *phonological representation* is used in this study somewhat loosely; it is intended to index the emerging property of the brain that represents, in an increasingly fine-grained and robust manner, the linguistic construct of a phoneme.

We acknowledge the fact that there are related lines of research in developmental dyslexia that currently investigate the intricate dynamics of auditory processing and speech perception of acoustic signals (e.g., Serniclaes, Van Heghe, Mousty, Carre, & Sprenger-Charolle, 2004). One obstacle in our field is that linguistic constructs such as the emergence or segmentation of a phonological representation, which cognitive psychologists and linguists are prone to study, need to be integrated with constructs such as auditory processing and phonetic categorization of the speech signal, which are more commonly studied by speech scientists. Although a speech perception task was used in the current study, it was not our intention to attempt to measure the relative contribution of categorical versus segmentation capacities in predicting reading and reading-related processes. Instead, our primary goal was to use three convergent cognitive tasks to infer that a deficit exists underneath that of PA, a weakness that does not invoke a metacognitive level of processing. The exact nature of the deficit in phonological representations is necessarily still vague, at this point, as most cognitive–linguistic constructs are. Further specification of the deficit at a linguistic level will be possible only after we understand and integrate the nature of bottom-up processes that impinge on the development of the neural networks that instantiate the phonological representation. By measuring the categorical capacities of participants concurrently, however, we hope to at least associate the two types of deficits, so that the relation between bottom-up processes of speech analysis and top-down processes of phonological–lexical activation can begin to be understood.

Speech and auditory perception studies have highlighted the differences among acoustic, phonetic, and phonemic levels of processing. Acoustic information in speech is a very complex waveform; nevertheless, certain invariances are abstracted by listeners, allowing for the emergence of categorical perception functions. In the latter, phonetic contrasts are greatest across a boundary, although some within-contrast variation is still apparent. With regard to the emergence of a phonemic representation, however, within-category (phonetic) contrasts are deemed irrelevant; they are treated as identical with respect to the linguistic phonemic category.

Studies investigating the nature of speech perception in participants with dyslexia, however, have yielded inconsistent evidence for speech perception difficulties. Interpretation of individual studies and generalization across studies have been hindered by differences in methodological approach and sampling issues as well as by theoretical considerations. Nevertheless, as we describe briefly in what follows of the Introduction, these types of studies have laid the foundation for the hypothesis that phonological processing is possibly deviant in individuals with dyslexia.

Studies investigating the effects of noise and lexicality on the perception of speech stimuli have yielded mixed results. For example, Brady, Shankweiler, and Mann (1983) showed that individuals with dyslexia were differentially affected by noise; however, these findings were not replicated by Snowling, Goulandris, Bowlby, and Howell (1986) or Pennington, Van Orden, Smith, Green, and Haith (1990), who found no group by noise interactions. Snowling (1981), Snowling et al. (1986), and Brady et al. (1983) showed that repetition of less familiar words and nonwords was less accurate in young children with dyslexia relative to reading age controls, but this was not replicated in adults with dyslexia (Pennington et al., 1990). More recently, a study by Gallagher, Frith, and Snowling (2000) found that children at risk for dyslexia were significantly worse at nonword repetition when the task used nonwords with unusual weak–strong stress patterns. Overall, researchers have interpreted these general perception and nonword repetition results as supporting the hypothesis that the phonological processing of participants with dyslexia is at least inconsistent, if not delayed or disordered, for speech stimuli.

Other types of speech perception studies involving identification and discrimination tasks have yielded differences between participants with dyslexia and controls. These tasks traditionally have used stimuli that varied along one or two acoustic dimensions. Some of the relevant findings include a higher proportion of participants with dyslexia not passing easier criterion items, less accuracy at the tails of the identification function, decreased slope of the identification function, and less predictability of the discrimination functions from the identification functions (Godfrey, Syrdal-Lasky, Millay, & Knox, 1981; Reed, 1989; Steffens, Eilers, Gross-Glen, & Jallad, 1992; Werker & Tees, 1987). In Godfrey et al.'s (1981) study in particular, participants with dyslexia performed worse than controls in discriminating synthetic stimuli along a /ba-da/ continuum than along a /da-ga/ continuum. This interaction, and the ability to predict each group's discrimination function from the stimulus identification data, suggested that the groups actually differed on perceptual ability and not just on general attention or extraneous task-specific cognitive factors. However, the fact that the categorical boundary did not differ from that of controls led the authors to suggest that the deficit may lie in the consistency of phonetic classification given a set of auditory cues rather than in the ability to perceive or discriminate these cues.

Although some categorical perception and discrimination results have been found to correlate with reading and PA ability (e.g., DeWeirdt, 1988; Godfrey et al., 1981; Nittrouer, 1999), the results of other studies have been less clear. At times, participants with dyslexia have demonstrated a shallower slope on the discrimination function, whereas other studies have reported group differences only at the ends of the continua where stimuli should be easier to categorize. In sum, it is difficult to glean from these earlier experiments what the exact nature of the perceptual difference might be in participants with dyslexia. For example, are the participants with dyslexia less able to perceive or use acoustic features to help define a phonemic category, or are they just more sensitive to within-boundary phonetic contrasts, thereby increasing the variability in their performance? Are the deficits found for a wide array of speech sounds? Can some of the findings be explained by general factors such as attention and IQ? To what extent are the differences between participants with dyslexia and controls just a reflection of differences in reading experience? These questions have resulted in further investigation of the relation between speech perception and reading ability.

There are additional potential reasons for the mixed results from speech perception studies. First, the acoustic features that typically are manipulated might not be the (only) relevant attributes of the acoustic signal that hinder the emergence of phonetic or phonemic distinctions. Second, the perception of phonetic contrasts does not necessarily imply that phonological representations are segmented effectively at the phonemic level. Thus, the representation may be robust enough to perform categorical perception experiments where one or two acoustic cues are systematically varied, but the representation might not be specified enough to efficiently perform PA and reading tasks at an appropriate level. Finally, the acoustic features that are relevant may vary across development, such that experimental manipulations may discriminate better at some ages than at others.

Nittrouer (1992) provided evidence for such an age-related phenomenon. Her study showed that children and adults differed in the extent to which they perceptually weigh dynamic versus static cues in the acoustic stream when making categorization decisions. Furthermore, the shift from the use of dynamic cues to that of static cues follows an age-dependent trajectory. Further experiments by Nittrouer (1996) have shown that the extent to which children show the more mature weighting pattern of acoustic cues is associated with linguistic experience. Chronic otitis media and low socioeconomic status (SES) both were related to immature weighting strategies and to poorer PA, even when nonverbal IQ was used as a covariate. One possible interpretation of these results vis-à-vis the development of phonological representations is that as the latter become more segmental, acoustic dimensions that covary best with phonemic structure receive greater weight.

Nittrouer's Developmental Weighting Shift hypothesis is also consistent with a computational model of phonological development implemented by Markey (1994), where the segmentation of the speech stream is an emergent property of the system. Thus, top-down demands and characteristics, together with articulatory output representations (and restrictions), influence how perceptual information is used and represented. Top-down demands include vocabulary acquisition, which Walley (1993) explained might increase the need to represent phonologically similar words at a phonemic level to avoid errors in an increasingly crowded lexical representational space. Researchers have hypothesized that the deficits in vocabulary acquisition in individuals with dyslexia may be related to poor phonological representations that do not allow efficient lexical–phonological encoding and retrieval to occur (Gathercole & Baddeley, 1990; Wolf & Obregon, 1992).

More recent studies in the literature have used different methods to investigate implicit phonological representations in typically developing individuals as well as in reading-impaired populations. Lance, Swanson, and Peterson (1997) attempted to validate an implicit phonological awareness task that was first proposed by Messer (1967). The task consisted of having participants identify which stimulus of a pair of nonsense words violated the rules of consonant combination in English. Participants needed to choose which stimulus from a pair sounded more like a real word (e.g., *shrib–shkib*). Results on this task were significantly correlated with results on explicit PA tasks, a multisyllabic word production task, and two reading outcome measures. Lance et al. (1997) indicated, however, that implicit-level tasks such as the one they used do not require segmental knowledge per se and thus may reflect a more holistic level of phonological awareness. It is possible that participants analyzed the characteristics of the stimuli as a whole and determined that certain acoustic combinations are not plausible or not part of a linguistic experience base. In this manner, phonemic-level representations might not have been necessary to do the task.

A study by de Gelder and Vroomen (1991) used a syllable similarity task to measure whether adults with dyslexia made errors based on consonant or syllabic structure of the stimuli. The primary measure was a paired associate learning task with novel nonwords. This task was implicit in that the dependent variable was derived from the types of recall errors produced. In addition, it required only a pointing response, reducing articulatory confounds. Because the de Gelder and Vroomen study demonstrated differences between adults with dyslexia and controls, a very similar task was chosen for the current study.

Matsala (1997) used a lexical gating paradigm to compare word recognition ability in children with dyslexia and chronological age-matched controls. She found that 9-year-olds with dyslexia needed more of the speech input than did their normally achieving peers. Matsala also investigated word frequency and neighborhood density effects on the children's ability to identify the target words. She documented a reading group by neighborhood density interaction, such that participants with dyslexia needed more speech input for words with sparse neighborhood densities than did controls. Word frequency affected both groups in a similar manner, with high-frequency words requiring less of the speech input than low-frequency words. Task performance on the gating paradigm was correlated with explicit PA, and single-word and pseudo-word reading tasks were administered concurrently.

The current study used a version of Matsala's (1997) lexical gating task but extended the research by accounting for potential IQ and reading experience confounds. In addition, responses to the first gated interval were coded for accurate identification of the initial phoneme, a measure that Metsala had used in a prior experiment with typically developing participants. Finally, the current study related the gating paradigm results to other measures of implicit phonological representation as well as to associated abilities that are commonly found to be deficient in children with dyslexia such as verbal short-term memory and rapid automatized naming (RAN).

The goal of the current study was to test the segmentation hypothesis of dyslexia by applying different methods, using a wide array of speech stimuli, to show that implicit phonological representations are deficient in dyslexia. The experimental methods used attempted to reduce metalinguistic and articulatory requirements as well as the more general confounds of attention, nonverbal IQ, and reading experience. We measured the relation between implicit phonological representation skills and speech perception, PA, literacy, and literacy-related abilities.

Because dyslexia often is found to be comorbid with articulation and language difficulties, we decided to investigate the current set of hypotheses regarding implicit phonological representations in two subsets of participants with dyslexia: those who have a history of speech sound (articulation) disorder (SSD) or language impairment (LI) and those who have no such history. It is plausible that children with an extended phenotype of dyslexia that includes early or ongoing articulatory and linguistic deficits are the ones for whom implicit phonological representation or speech perception deficits are more apparent (Harm & Seidenberg, 1999; Manis et al., 1997).

Specific hypotheses

Hypothesis 1. As predicted by the segmentation hypothesis, participants with dyslexia will perform worse on three measures of implicit phonological representation (a syllable similarity task, a gated lexical decision task, and a priming task) compared with both chronological age- and reading age-matched controls.

Hypothesis 2. Participants' poorer performance on implicit phonological representation tasks will be correlated with a less mature perceptual weighting pattern on a speech perception paradigm (Nittrouer, 1992).

Hypothesis 3. Measures of implicit phonological representation from the three experimental tasks will be positively correlated with measures of PA ability, short-term phonological memory, rapid lexical retrieval, and reading and spelling skills.

Hypothesis 4. To the extent that articulatory and oral language deficits are markers for a broader and possibly more severe deficit in phonological processing, children with dyslexia with these comorbid disorders will demonstrate greater deficits in speech perception and implicit phonological representations.

Method

General rationale

Given the inherent difficulty of measuring implicit cognitive processes, this study included three experimental tasks for this purpose: a syllable similarity task, a gated lexical retrieval task, and a priming task. These three tasks were quite distinct in their surface characteristics but attempted to measure the same underlying construct. In this manner, convergent validity could be established for the construct while at the same time reducing the chance that significant results are due to task-specific factors.

Participants

The following four groups of 20 participants each were included in this study: two groups of 11- to 13-year-olds with dyslexia (one with, and one without, a history of speech sound disorder [SSD] and/or specific language impairment [SLI]), a chronological age-matched control group, and a reading age-matched control group. Participants for the two groups with dyslexia (hereafter labeled RD and RD+ groups) and for the chronological age-matched control group (CA group) were recruited from a database of twins participating in a larger study conducted by the Colorado Learning Disabilities Research Center (CLDRC) (DeFries et al., 1997). Ascertainment methods and demographic characteristics for the original twin sample are described in detail in that study. The original twin sample is large and ethnically heterogeneous, facilitating selection of ethnic minorities in accordance with their population base rates in Colorado.

For the current study, twin pairs currently between 11 and 13 years of age with at least one twin identified as having a reading difficulty were selected initially. Because all of these twins had previously been tested on a comprehensive reading battery as part of the larger twin project, a reading composite score was available for each of them. This reading composite score was computed using an algorithm from a discriminant function analysis used to identify children in a previous nontwin sample with a history of dyslexia (DeFries, 1985). This analysis incorporated three subtests from the Peabody Individual Achievement Test (PIAT) weighted in the following manner: Reading Recognition, 50%; Reading Comprehension, 25%; and Spelling, 25%. The composite score is scaled so that a score less than zero is indicative of a reading disability. Twin pairs chosen as potential participants for the current study had a school history of reading difficulty and also a negative discriminant function composite score. If both twins within a twin pair met these criteria, one was selected at random for inclusion in the current study.

Children with dyslexia were further categorized as having a history of speech and language difficulties or not based on a speech and language parent questionnaire. Parents supplied information regarding each participant's speech and language history, including any past or present difficulties with articulation, morphology, syntax, word finding, semantic development, and language comprehension. In addition, parents were asked whether their children ever were identified as having an articulation and/or language disorder by a speech–language pathologist and whether they had received treatment. Affirmative responses to the latter two questions led to the children being placed in the RD+ group. Children with no articulation or language difficulties were placed in the RD group.

Chronological age-matched controls were also selected from the larger twin sample. All twin pairs currently 11–13 years of age who did not have a history of reading problems at school were selected initially. Both twins in each pair needed to have a reading discriminant function score greater than zero to remain in the group of potential participants. One twin from each pair was selected using a stratified random sampling method so as to match groups on chronological age, gender, SES, and Performance IQ (PIQ). Thus, a specific twin from a twin pair may have been selected to equalize the groups on one of the other matching variables if both met the initial reading criteria. Groups were matched on PIQ, rather than Verbal IQ or Full Scale IQ, to reduce the effect of regression to different means in statistical analyses.

The reading age-matched control group (RA) could not be obtained from the larger CLDRC twin sample because there were no control twin pairs currently 8 years of age in the sample. Instead, the RA control group was selected from a developmental participant pool list maintained by the psychology department at the University of Denver. A current administration of the PIAT Reading Recognition subtest yielded a raw score for single-word reading that was used to match the RA controls to the children in the groups with dyslexia. Because the RD+ group obtained a slightly lower PIAT Reading Recognition raw score, the RA control group was selected so as to match to the more severe of the two groups with dyslexia (making group members slightly younger than they otherwise might have been). In addition, the RA controls were matched to the other groups on gender, SES, and PIQ.

All participants were monolingual speakers of English and had no known neurological compromise. Children were excluded if they had a PIQ less than 80. All children passed a pure tone hearing screen at 0.5, 1, 2, and 4 kHz bilaterally at 25 dBHL at the time of testing. Participants were selected regardless of inattentive or hyperactive symptoms because attention deficit/hyperactivity disorder (ADHD) is known to be comorbid with dyslexia. ADHD symptoms were measured and used as a covariate in analyses when necessary. The Hollingshead Four-Factor Index of Social Status was used as a measure of SES.

The proposed research did not take advantage of the twin status (i.e., only one twin from a given pair participated), but it did make use of the extensive cognitive, achievement, and psychiatric testing information that already has been accumulated on these children. In addition to the PIAT-Revised (PIAT-R) reading battery, these children were given the Wechsler Intelligence Scales for Children-Revised (WISC-R), from which the PIQ information was obtained. RA control participants, for whom this information was not available, were given the three PIAT-R subtests to calculate a current reading discriminant function score. In addition, they were administered the Matrices and Block Design subtests of the Wechsler Abbreviated Scale of Intelligence (WASI), which provides a reliable and valid PIQ estimate.

All children were asked to participate via a letter sent to their home or via a telephone call. All 11- to 13-year-old participants were tested in one $2\frac{1}{2}$-h session in a laboratory setting. They were given a break at the midpoint of the testing session. The 8-year-old (RA) participants were tested over two sessions, each lasting approximately $1\frac{1}{2}$ h. All children were paid for their participation.

Measures used for group selection and matching

Age, parental education and occupation, and gender information was obtained from a demographic questionnaire completed by the parents. All children were currently administered the PIAT-R Reading Recognition subtest to select the reading age-matched control group. Because many of the participants with dyslexia received, or were currently receiving, intervention for their reading deficits, no decisions regarding inclusion were based on current performance on the Reading Recognition measure. It is assumed that the participants with dyslexia have benefited from intervention and that the severity of their deficits may be less now than in the past. As a result, this sample provided a conservative test of the hypotheses of this study because the impaired groups may have partly compensated by this time.

Experimental tasks

Syllable similarity task

This task was adopted and modified from previous research by Treiman and Breaux (1982), where differences were documented between children and adults in the way they used either phonemic or syllabic structure to identify which two of three spoken syllables were similar. Stimulus triads in Treiman and Breaux's study consisted of three words, two of which had the same initial phoneme and two of which had similar syllabic structure (e.g., *bis*, *dis*, *bun*). These triads were used in a paired associate learning and memory task, and confusion errors were tallied for learning trials as well as for a delay trial. This task did not require explicit judgments of the similarity between syllables; rather, the implicit categorizations emerged in participants' errors (i.e., making a syllabic similarity error, as compared with a common phoneme error, when recalling the names of the animals). Treiman and Breaux found that children used syllabic structure more than phonemic structure during the learning phase than did adults, whereas both children and adults used phonemic similarity to a greater extent during delayed recall.

The stimuli from Treiman and Breaux (1982) were audiotaped. Five unique triads were used as trial items, with the same relations between syllables as described previously. A total of 15 different small toy animals, all of different species (e.g., horse, cow, and elephant), were used in the five triads. No animal in a given triad had a species name that began with the same phoneme as any of the novel syllable names to be learned. After children had demonstrated that they knew the species name of each animal, they were was told that they needed to learn the "made-up names" of the toy animals. Via headphones, the audiotape provided the names of the three animals (e.g., "The horse is /diz/. The cow is

/bun/. The elephant is /bis/.''). On the first trial, each child was asked to repeat the syllable names, and mispronunciations were corrected. The child was then asked, via headphones, what animal corresponded to each of the syllable names (e.g., ''Which one is /diz/?'' ''Which one is /bis/?'' etc.). The child responded by pointing to the correct animal. Teaching and testing trials alternated until six learning trials had elapsed or the child had two successive errorless test trials. If the child reached criterion, a delay of 2 min was introduced, during which the child named letters or numbers from various printed distractor stimuli cards. After the delay, the child was asked the names of the animals once again as in previous test trials (but without a training trial preceding it). If all animals were named correctly, the delay interval was lengthened by 30 s for the next triad. The dependent variables were the number of confusions that were either consonant errors, syllabic errors, or anomalous errors.

Each confusion error that could be assigned unambiguously (one correct and two transposed) was coded as a consonant error, a syllabic error, or an anomalous error. For example, if the child pointed to /diz/ correctly but transposed /bis/ and /bun/, it was counted as a consonant error (i.e., the child encoded the syllable names using information about the first phoneme, among other features, and was now confusing the two at this level). If the child correctly identified /bun/ but confused /bis/ and /diz/, it was coded as a syllabic error (i.e., the child encoded the syllable name at the level of syllabic structure and confused the two names at that level). If the child made any other combination of errors, it was coded as anomalous.

Lexical gating task

This task is one initially developed by Grosjean (1980) and more recently used by Matsala (1997) and Dollaghan (1998) to measure word recognition given increasingly larger acoustic segments of a word. The first segment (or *gate*) presented was 120 ms in length, with subsequent segments adding 60 ms to the previous segment. Grosjean's (1980) study showed that 5- to 7-year-olds required more acoustic input to recognize words than did adults. Dollaghan (1998) used the same paradigm to measure deficits in lexical access in children with SLI but selected target words carefully to control for familiarity effects. Participants with SLI required significantly more of the acoustic signal than did their peers to identify the unfamiliar words but not the familiar words. More relevant to this study was the additional finding that participants with SLI were significantly less likely to respond with correct initial consonants at the earliest gated interval than were their peers for all word types. Dollaghan concluded that both representational and auditory–perceptual inefficiencies of the phonological system may be contributing to slowed lexical access in children with SLI. Thus, this deficit in initial consonant accuracy may be due partly to poor phonemic segmentation and representation, without which the segment of the word at the initial gated interval cannot be identified.

In the current study, Hypothesis 1 predicts that participants with dyslexia will be less able than controls to identify the correct initial phoneme of the target word at the earliest gated interval. Only highly familiar words were used to avoid a lexical learning efficiency confound. Thus, deficits in initial phoneme identification would be less likely to be due to poor access to a newly learned lexical item and more likely to be due to deficiency in linking acoustic information to a phonemic representation, which would then be associated with a lexical entry.

All of the target words, in addition to being high frequency, were nouns with a relatively low number of phonologically similar neighbors (i.e., neighborhood density was minimized). Words with a voiceless stop consonant as the initial phoneme were excluded. Five of Dollaghan's (1998) target words met all of these criteria and were used in this task. An additional 3 words were selected from Matsala's (1997) word list, with the remainder selected from the word lists provided by Rescorla (1989) and Reznick and Goldsmith (1989) said to characterize the expressive lexicons of 3-year-olds. A total of 16 words were selected. Each word was from a different onset–vowel phonological neighborhood containing anywhere between 1 and 10 phonological neighbors ($M = 5.56$, $SD = 2.70$). High word frequency status was corroborated using word lists compiled by Kucera and Francis (1967). Of the 16 words, 8 (*boat*, *bead*, *bus*, *dish*, *dirt*, *dog*, *girl*, and *game*) were selected to have an initial voiced stop consonant, whereas the remaining 8 were composed of two each from the following initial consonant categories: glide (*watch* and *wood*), nasal (*moose* and *mouth*), liquid (*light* and *rain*), and fricative (*feet* and *fork*). There were no differences between the 8 words with stop initial consonants and the 8 words with nonstop initial consonants on word frequency, $t < 1$, or number of phonological neighbors, $t < 1$. In addition, 3 other words were selected and prepared as practice items.

The 19 words (16 test items and 3 practice items) were digitally prepared and manipulated by John Hansen in his laboratory at the Center for Spoken Language Research at the University of Colorado–Boulder.[1] All segments for each word were presented to participants via computer using E-Prime software, with the 3 practice words presented first and followed by the 16 target words in random order. For each target word, the series of gates were given in the context of a guessing game. Each participant was told that increasingly longer fragments of words would be presented via headphones from the computer and that he or she should guess the whole word after each segment. Responses were captured via microphone, allowing the computer to measure reaction time (RT). Although RT was not used as a dependent measure in the current task, this method of administration provided consistency with the priming task and allowed the next stimulus item to be administered by E-Prime. The total number of word guesses containing the correct initial phoneme following presentation of the 120-ms gate was recorded. In addition, the respective gates at which the participant correctly identified the correct first phoneme, first syllable, and entire word were recorded for secondary analyses.

Priming task

This measure is somewhat analogous to the gating task except that word segments of different lengths (i.e., gated segments) were used to *prime* the target word. In the previous

---

[1] Target words were recorded on a Sony TCD-D8 digital audiotape recorder at a sample rate of 48 kHz, monoaurally, with an Audio-Technica microphone in a soundproof recording/listening chamber. Digitized data were transferred to a PC at a sample rate of 44.1 kHz using a .WAV format. Each word was hand-parsed, resulting in nine segments, the first 120 ms in duration and each subsequent segment adding 60 ms. All segment files were down-sampled to 22.5-kHz .WAV files. All of the original word tokens were measured for duration in milliseconds. The average word duration was 560 ms ($SD = 8$, range = 504–600). For the small proportion of target words whose duration was less than 540 ms, the last stimuli in the series was a repetition of the 540-ms segment with an additional 60 ms of silence added to the end. Otherwise, the last segment of the target word consisted of the remaining part of the word plus silence until the 600-ms duration was reached. In the end, this was inconsequential because all of these words were identified prior to the final segment for all participants.

task, participants listened to gated segments and guessed the target word; in this task, short segments primed the identification of the target word. Primes consisted of the first 120- or 240-ms segments of the target word. It was hypothesized that only a more robust, finely discriminated, and segmented phonological representation would be able to make use of the 120-ms prime to activate a phonological representation. This activation would, in turn, facilitate the identification and retrieval of the target word that followed. A participant whose phonological representation was not fully or adequately segmented would not make adequate use of the 120-ms prime to activate the phonological–lexical network. The 240-ms duration prime was hypothesized to facilitate retrieval for all participants because enough acoustic information would be provided to activate even the more deficient (poorly segmented) phonological representation.

A total of 48 words were selected as targets for the priming task. They all were highly familiar (in the expressive repertoire of 3- and 4-year-olds) high-frequency nouns, with relatively sparse phonological neighborhoods. None of the words overlapped with those chosen for the lexical gating task, but these words were chosen from the same vocabulary lists (for a list of the stimuli, see the Appendix). All words had different onset–vowel combinations, with 24 of the words beginning with a voiced stop consonant and the remaining 24 words chosen from the following consonant categories in approximately equal numbers: glides, liquids, nasals, fricatives, and affricates. The words on the list had an average of 7.02 ($SD = 2.00$) phonologically similar neighbors and an average frequency rating of 67.8 ($SD = 75.1$). A word was counted as a neighbor if it differed from the target word by one phoneme (by addition, omission, or substitution). The stop consonant words did not differ from the nonstop consonant words with regard to frequency, $t < 1$, or neighborhood density, $t < 1$. An additional 8 words were selected from the word lists as practice items.

Target words were recorded and digitized in the same manner as described in the lexical gating task. Four different types of primes were generated for each word. Two segments were spliced from the front part of each word: one segment that was 120 ms in duration and the other that was 240 ms in duration. The third type of prime was a digitally stretched version of the 120-ms prime, which is described here only because it was part of the Latin square design but is beyond the scope of this study. The fourth prime type was a neutral prime, consisting of a nonspeech complex tone, modeled after "Tone 1" in the Tallal and Percy (1974) study. Four versions of the priming task were generated, with the four types of primes paired with each subset of 12 words in a Latin square design. Thus, the first 12 words of the list were preceded by a neutral prime in Version 1, by a 120 ms prime in Version 2, by a stretched prime in Version 3, and by a 240 ms prime in Version 4. This design was adopted to minimize any lexical confound, whereby certain prime–target pairings may be easier to process than others. This counterbalancing of versions across participants in each group also corrected for any slight duration differences in the target words themselves that otherwise would have affected the response time data. Each subset of 12 words contained 6 words with an initial stop consonant, 2 words with an initial fricative, and 1 word each with an initial glide, a nasal, a liquid, and an affricate.

To prevent ceiling effects in the priming data, the target words were degraded with a White-Gaussian noise mask. The noise was generated by a random noise generator at an equivalent sampling rate. The process was automated, such that the target words were degraded with the White-Gaussian noise at different signal/noise (SNR) levels. Pilot testing of the noise-masked words indicated that a -10 SNR produced a 60–70% correct identification rate in control pilot participants, sufficient to avoid ceiling effects.

The primes and noise-degraded target items were presented, appropriately paired via computer using the E-Prime software package. The program was designed to initially present 8 practice target words, encompassing all of the different types of prime. Subsequently, the 48 words (four sets of 12 words, each set paired with a different type of prime) were presented to each participant in a random order, with the only constraint being that no more than 2 words with the same prime type could be administered consecutively. The entire list of 48 words was presented twice for a total of 96 trials (Priming I) using standardized instructions.[2]

Participants were administered a second priming task (Priming II) at the end of the testing session (or on the second testing session for 8-year-olds), whereby the same 48 target words were paired with mismatched primes (i.e., segments from one word were used as primes for a different word). The same four types of primes and counterbalanced prime–target pairing method were used. This resulted in four different versions of the mismatched priming task, counterbalanced across participants within each group. The 48 words were again presented in pseudorandom order but were not repeated, resulting in a total of 48 trials.

Four dependent variables were obtained from each priming task (I and II). The number of correctly identified words for each prime type was tallied (24 possible per prime type in Priming I, 12 possible per prime type in Priming II). In addition, the number of trials where the participant identified the first phoneme correctly (regardless of the accuracy of the resulting word) was also recorded. The average response time for each of these two sets of correct items (Word RT and Letter RT) was also recorded.

It was hypothesized that in the congruent condition (Priming I), CA and RA participants would demonstrate an increased number of correctly identified words and initial phonemes in the 120- and 240-ms prime type conditions relative to the neutral condition. Participants with dyslexia would demonstrate facilitation only when the 240-ms prime was used. In the incongruent condition (Priming II), an inverse pattern was expected, whereby CA and RA participants would show interference in both accuracy and latency for targets preceded by 120- and 240-ms primes, whereas participants with dyslexia would demonstrate interference only with the 240-ms prime.

Reading component measures

PA and phonological coding constructs, both component reading processes closely associated with reading ability, were measured. The phoneme reversal task and the Pig Latin task, measuring PA, were administered as described by Pennington et al.

---

[2] Participants were told that they would hear a word masked with noise via headphones. They were instructed to identify the noise-masked word as quickly as possible. They were alerted to the fact that a "sound" would precede each word but were told to focus on identifying the subsequent noise-masked word. Response time was recorded via microphone, which acted as an external response box. The order of events, as perceived by the child, was as follows. First, a screen appeared with the word "Ready?" Second, a blue screen with a centered green cross appeared and remained on the screen until the child's verbal response was recorded. Auditorily presented prime lagged the onset of the blue screen by 1 s. Third, the "Ready?" screen reappeared until the onset of the next trial. An interstimulus interval of 300 ms was used between the prime offset and the target word onset. The response time clock began at the onset of the target word and was stopped by the child's response. The examiner controlled the onset of each new trial.

(1990). In the former, the child heard a monosyllable word and was asked to reverse the sounds to form another word (e.g., *lace* becomes *sail*). Stimuli included practice items and 24 trial items. The number of correctly produced reversed words was the dependent variable.

The Pig Latin task, containing 26 real words as items, required each child to take the first phoneme and place it at the end of the word followed by *ay* (e.g., *go–ogay, man–anmay, drip–ripday*). Mono- and disyllabic words were used, some with singleton word onsets and some with word-initial clusters. The number of correctly manipulated items was used as the dependent variable.

The Word Attack subtest of the Woodcock Johnson Tests of Achievement-Revised was administered as a measure of phonological coding. Standard administration procedures were followed, with the exception that all words on the list were administered. This measure primarily requires the child to use sound–letter correspondence rules to decode pseudowords, although some items also require knowledge of the exception and orthographic rules of English. The number of correctly pronounced nonwords, out of a total of 30, was used as the dependent variable.

Other measures associated with reading ability and disability

Dollaghan and Campbell's (1998) nonword repetition task was administered as a measure of phonological short-term memory. Their measure was carefully constructed to avoid syllable lexicality and phoneme predictability confounds and has shown an inverse relation between nonword length and repetition accuracy in a group of children with SLI. Dollaghan and Campbell's stimuli were presented via audiotape, and responses were recorded. Repetition accuracy was scored in terms of the total number of correct phonemes articulated. A portion of the audiotaped responses was transcribed by a licensed speech–language pathologist blind to group status for reliability purposes.

The Color Naming and Object Naming subtests of the Comprehensive Test of Phonological Processing were used as measures of rapid serial naming. Color Naming and Object Naming are continuous naming tasks, much like the letter or number naming tasks that have differentiated participants with dyslexia in prior studies (Denckla & Rudel, 1976; Katz, 1986). The alphanumeric versions of this task were avoided because performance could be confounded by differential reading and academic experience in participants with dyslexia. The total time (in seconds) for naming the stimuli in Parts I and II of each condition was used as the dependent variable.

Speech perception task

A speech perception identification task, obtained from Susan Nittrouer, was administered to all participants. The stimuli independently manipulated both fricative noise and vowel transition in a factorial design. Nine synthetic noises across the /s/ to /sh/ continuum (2.2–3.8 kHz in 200-Hz intervals) were paired with naturally produced vocalic portions with formant transitions appropriate for a preceding /s/ or /sh/. These transitions and vocalic portions were produced by removing the preceding fricative noise from a male speaker saying /sha/, /sa/, /shu/, and /su/. The participants were asked to label the synthesized syllable they heard by saying that syllable out loud and then pointing to the

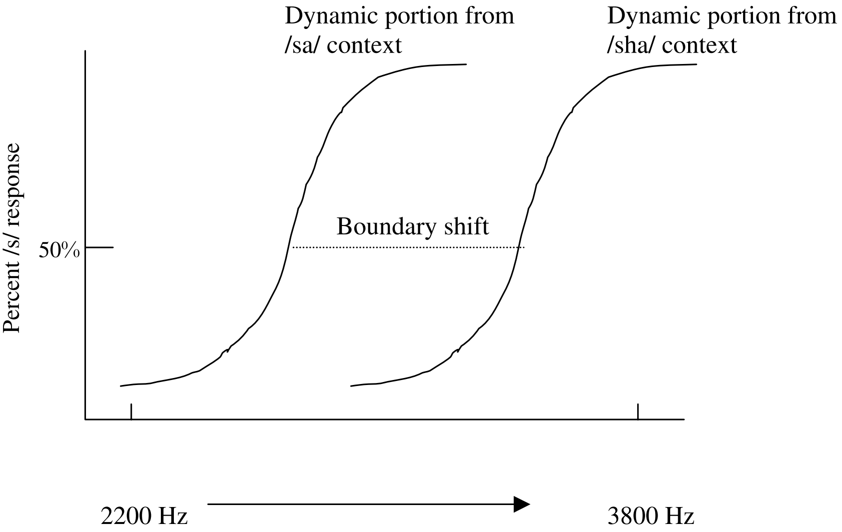

Fig. 1. Mature categorical perception pattern in speech perception labeling task. The slope of each curve yields a datum point, as does the distance between the two curves where the percentage of /s/ responses is at chance (vowel context distance). The slope of each line indicates how sensitive the listener is to changes in frequency. The distance between functions indicates sensitivity to the dynamic portion of the consonant–vowel transition.

appropriate picture.[3] The proportion of /s/ responses in each stimulus condition underwent a probit transformation, and data analysis was conducted in accordance with Nittrouer's (1992) analysis procedure. Fig. 1 shows a mature categorical response pattern.

Results

Preliminary analyses

Prior to conducting inferential statistics, all dependent variables and potential covariates were checked for skew, kurtosis, and outliers. Nearly all variables met criteria for normality. For those few variables that did not, one outlier was identified in each instance that caused the significant skewness. The outlier's score was changed so that it would be 2.5 *SD* from the mean (and beyond all other data points). This procedure was preferred so as to maintain equal group sample sizes. All distributions became normally distributed after this correction was implemented. Homogeneity of variance tests was also conducted during each analysis, and these showed few violations. In the cases where there was a violation,

---

[3] Children were told that they would be listening to one of two names—/sa/ (a space creature) or /sha/ (a king from a far-away land); /su/ (a girl) or /shu/ (a shoe)—spoken by a computer via headphones. After 5 live voice practice trials, children were given 10 practice trials with stimuli at ends of the fricative continuum, where they needed to reach 90% criterion to proceed to the test trials. Actual trials, consisting of five sets of 18 stimuli (9 points along the fricative dimension × 2 transition contexts [/s/ and /sh/] for the /a/ vowel), followed the practice trials. The entire sequence was repeated for the /u/ vowel condition. Vowel order was counterbalanced across participants in each group. The computer recorded the responses via a button-press response box with pictures representing each syllable next to each button.

the ratio of the largest group variance to the smallest group variance never was greater than 3.0. Even so, significance tests for specific contrasts, or post hoc comparisons, were reported with Welch's correction for unequal variances. In cases where the sphericity assumption was violated in a repeated-measures analysis of variance (ANOVA), statistical results for the within-participants and interaction effects were reported using the Greenhouse–Geisser degrees of freedom correction. The alpha level used for all analyses was .05.

Analyses of demographic, selection, and matching variables

The sample as a whole was ethnically diverse and was approximately representative of the ethnic composition of the Denver metropolitan area. The ethnic breakdown of this sample was as follows: 78.8% Caucasian, 3.8% American Indian/Alaska Native, 1.3% African American, 3.8% Asian or Pacific Islander, 6.3% Hispanic, and 5.0% "other." The gender distributions across groups were also quite similar, with 40% females in the RD groups and 50% females in the control groups.

As can be seen in Table 1, the RD, RD+, and CA control groups were of similar chronological age (overall range 132–170 months of age), whereas the RA control group was significantly younger (96–106 months of age). Although the overall one-way ANOVA was significant, $F(3, 76) = 136.60$, $p < .001$, post hoc comparisons revealed that the only difference among the groups was the younger age of the RA controls. Omnibus tests from similar one-way ANOVAs computed on maternal and paternal years of education were not significant at the $p < .05$ level; however, a trend was observed in both cases. Post hoc comparisons showed that maternal and paternal education tended to be higher in the CA and RA control groups. This was not an unexpected finding given that parents of children with dyslexia often have reading difficulties themselves, which may reduce their overall academic achievement. However, when occupation was taken into consideration, as it is in Hollingshead's Four Factor Index of Social Status, there were no group differences, $F < 1$. The latter indicated that the families that participated in this study ranged from middle- to upper middle-class segments of the population.

The RD and RD+ groups were classified as reading disabled based on the reading composite score computed via the discriminant function algorithm used in the larger twin study conducted by the CLDRC. Although the RD+ population, being comorbid for a history of speech and/or language impairment, could be hypothesized to have a greater severity of impairment, the two dyslexia groups in this study were matched on the reading composite score to avoid a severity confound when analyzing the results on the experimental measures. Although the omnibus ANOVA shows a significant difference among the four groups, $F(3, 76) = 86.60$, $p < .001$, post hoc comparisons revealed that there was no difference between the two groups with dyslexia. Likewise, there was no significant difference between the CA and RA control groups.

All children were administered the PIAT-R Reading Recognition subtest to match the RA control group to the groups with dyslexia on reading experience. As expected, the omnibus ANOVA was significant, $F(3, 76) = 45.60$, $p < .001$, with the CA controls having a higher score than all of the other groups. In addition, the RD+ group scored significantly worse than the RD group, which may be due partly to a regression artifact. In other words, the RD and RD+ groups, although matched on reading composite score in this study, could have regressed to their own population means on the PIAT-R Reading Recognition subtest, alluding to the probability that the RD+ population has a more severe

Table 1
Demographic and group selection variable means (and standard deviations) by group

Variable	RD	RD + SSD/SLI	CA control	RA control	F value	Significance
Demographic variables						
Age (months) and range	147.4 (9.4)$_a$ 132–165	147.7 (10.9)$_a$ 132–170	144.7 (9.4)$_a$ 132–163	101.0 (3.1)$_b$ 96–106	136.61	$p < .001$
Gender	40% female	40% female	50% female	50% female	—	—
Father education	14.9 (3.5)$_a$	14.3 (2.4)$_a$	16.1 (1.6)$_a$	16.2 (1.7)$_a$	2.71	$p = .06$
Mother education	15.4 (2.8)$_a$	14.3 (2.3)$_a$	16.0 (2.3)$_a$	16.1 (2.1)$_a$	2.45	$p = .07$
SES	47.9 (11.0)$_a$	45.1 (10.9)$_a$	46.5 (11.2)$_a$	47.5 (10.4)$_a$	0.89	$p = .45$
Selection and matching variables						
PIAT Reading Recognition	45.7 (7.3)$_a$	39.6 (7.2)$_b$	61.8 (7.0)$_c$	40.7 (5.8)$_{ab}$	45.64	$p < .001$
PIQ	102.4 (10.4)$_a$	100.9 (11.9)$_a$	105.6 (6.9)$_a$	108.1 (11.1)$_a$	1.99	$p = .12$
Reading composite	−1.01 (0.65)$_a$	−1.18 (0.81)$_a$	1.81 (0.88)$_b$	1.46 (0.71)$_b$	86.57	$p < .001$
ADHD composite	13.1 (10.5)$_a$	16.4 (11.8)$_a$	8.3 (5.3)$_a$	10.4 (8.3)$_a$	2.54	$p < .07$

Note. A one-way ANOVA was computed on all variables ($df = 3, 76$). $N = 20$ for all groups. Standard deviations are in parentheses. Group means with no common subscripts are significantly different at the $p < .05$ level using the Tukey procedure.

reading problem, in general, than the RD population. Finally, the RA control mean ($M = 40.7$, $SD = 5.8$) was in between that of the RD mean ($M = 45.7$, $SD = 7.3$) and RD+ mean ($M = 39.6$, $SD = 7.2$) and was not significantly different from either of them.

Parents were asked to complete the ADHD Rating Scale, which consists of the 18 DSM-IV symptom criteria (American Psychiatric Association, 1994) transformed into a Likert scale (range 0–3). The sum of the responses for each item was used as the ADHD composite score. The omnibus ANOVA revealed a trend for group differences, $F(3, 76) = 2.54$, $p = .07$, with both the RD and RD+ groups (*M*s = 13.1 and 16.4, respectively) obtaining higher total symptom scores than the CA and RA control groups (*M*s = 8.3 and 10.4, respectively). Likewise, it is known from previous research that children with dyslexia have slightly lower Full Scale and Verbal IQ scores than do controls. In an effort to avoid differential regression, the groups were matched on PIQ from the WISC-R (WASI for RA controls). Although the omnibus ANOVA revealed no significant differences, $F(3, 76) = 1.99$, $p = .12$, the expected trend of slightly lower PIQ for participants with dyslexia relative to controls was evident. Due to these trends, both ADHD symptoms and PIQ were used as covariates in follow-up analyses when the covariate was significantly correlated with the relevant dependent variable.

Experimental tasks

The analyses of the three experimental tasks are organized similarly. First, data reduction procedures are described. Next, the RD and RD+ groups were compared to see whether there were any significant differences along the relevant dimensions. If no differences existed and there were no interactions of group with a within-participants variable, the RD and RD+ groups were collapsed into one group. Subsequently, the combined RD group was compared with the CA control group using the appropriate analysis, with contrasts and post hoc comparisons performed as appropriate. The combined RD group was then compared with the RA group in a similar manner. Any secondary analyses for each experimental task are reported last.

Analysis of the syllabic similarity task

The administration procedures of this task dictated that children be given subsequent learning trials until they achieved two correct trials in succession. Therefore, different numbers of total learning trials were possible across participants. To have data that could be compared across participants, the proportion of consonant and syllabic confusions relative to the total number of learning trials administered was computed for all four groups. The same proportions were calculated for the delay trial data. When comparing the RD, RD+, and CA controls, no main effect of time (learning vs. delay) or Group × Time interaction was evident in a Group (3) × Error Type (2) × Time (2) repeated-measures ANOVA (Time: $F(1, 57) = 2.19$, $p = .16$; Time × Group: $F(2, 57) = 0.46$, $p = .64$); thus, the learning and delay trial data were collapsed. New proportions were calculated with the number of total trials (learning + delay) in the denominator for each type of confusion error. Because the comparison of interest is the relative proportion of consonant errors to syllabic errors in each of the groups, a proportion difference score was calculated. The proportion of syllabic errors was subtracted from the proportion of consonant errors, resulting in a proportion difference score (Table 2).

Table 2
Proportions of consonant and syllable errors by group for the syllable similarity task

	Proportion of consonant errors	Proportion of syllable errors	Resulting difference score
Learning trials			
RD	.05 (.06)	.08 (.05)	−.03 (.08)
RD+	.03 (.04)	.11 (.08)	−.08 (.09)
CA controls	.06 (.06)	.03 (.05)	.03 (.12)
RA controls	.06 (.05)	.06 (.08)	−.00 (.09)[a]
Delay trials			
RD	.02 (.06)	.13 (.16)	−.11 (.18)
RD+	.06 (.15)	.16 (.15)	−.10 (.27)
CA controls	.11 (.14)	.02 (.06)	.09 (.15)
RA controls	.11 (.14)	.01 (.04)	.10 (.09)

Note. Proportions are out of total learning or delay trials administered. Standard deviations are in parentheses.
[a] Nonrounded C–S score in this cell is −.003.

Results of a one-way ANOVA on the proportion difference score comparing the RD, RD+, and CA groups revealed a significant omnibus test, $F(2, 57) = 9.11$, $p < .001$. Contrasts, selected a priori, compared the RD group with the RD+ group and compared each of the groups with dyslexia with the CA control group. There was no difference between the groups with dyslexia on the proportion difference score; both groups made approximately equal numbers of consonant and syllabic confusion errors. However, both the RD and RD+ groups had significantly lower proportion difference scores than did the CA control group (RD: $t(57) = 3.12$, $p < .01$; RD+, $t(57) = 4.08$, $p < .001$). Thus, both groups with dyslexia made significantly more syllabic confusions than consonant confusions.

Because there was no difference between the RD and RD+ groups, their data were combined into a combined RD group. A one-way ANOVA comparing the combined RD group with the CA control group on the proportion difference score was highly significant, $F(1, 52) = 18.08$, $p < .001$. Fig. 2 depicts the proportion difference scores for the combined RD group versus the CA control group. This finding supports Hypothesis 1, whereby participants with dyslexia used syllabic structure more than phonemic structure to encode the new animal names; this encoding pattern is reflected in the confusion errors they made on recall.

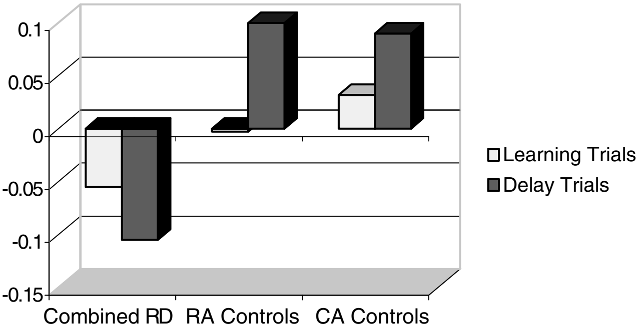

Fig. 2. Mean proportion differences (consonant errors–syllable errors) by group for the learning and delay trials of the syllable similarity task.

The RD, RD+, and RA controls were compared next. The Group (3) × Error Type (2) × Time (2) repeated-measures ANOVA using the proportional data for consonant and syllable errors separately revealed no main effect of time, $F < 1$, but there was a significant Group × Error Type × Time interaction, $F(4, 114) = 2.52$, $p < .05$. This significant result precluded collapsing the learning and delay trial data. Without knowing whether both RD groups were contributing to the interaction equally, they were not collapsed for the next analysis. The proportion difference score was computed separately for learning and delay trials. A Group (3) × Time (2) repeated-measures ANOVA revealed a main effect of group, $F(2, 57) = 6.67$, $p < .01$, and a Group × Time interaction, $F(2, 57) = 3.96$, $p < .05$. As can also be seen in Fig. 2, the RA controls had an equal number of consonant and syllable confusion errors during the error trials (C–S = −.003); however, after the delay, the RA controls made more consonant errors than syllabic errors. In contrast, not only did the RD and RD+ groups make slightly more syllabic errors than consonant errors during the learning trials, but this pattern also solidified over the delay condition, with the proportion difference score becoming even more negative. To determine whether there were any differences between the RD groups, they were compared directly in a 2 × 2 repeated-measures ANOVA. Their Group × Time interaction was not significant, $F < 1$, indicating that the interaction in the omnibus analysis was driven primarily by the divergence of the groups with dyslexia from the RA controls over the delay condition.

Comparing the CA and RA analyses, one can interpret the pattern of results as follows. The RA controls appear to consolidate their encoding process over the delay, with a more mature pattern of confusion errors emerging during the delay trials. The CA controls, in contrast, demonstrate the more mature pattern (a greater proportion of consonant errors than syllabic errors) from the beginning and thus do not demonstrate this change over time. Both groups with dyslexia demonstrate an immature pattern (a greater proportion of syllabic errors than consonant errors) during the learning trials and continue to demonstrate this pattern over the delay.

Analysis of the lexical gating task

The most relevant dependent measure from the lexical gating task with respect to Hypothesis 1 is the total number of correct first phoneme identifications at the earliest gate (120-ms segment). It should be noted that this task also had a stretched first gate added to the stimuli series for all participants, although this is beyond the scope of this article; statistical analyses include the stretched segment because it was part of the stimuli presented to the participants, but only the regular 120-ms gate data are discussed here.

The RD and RD+ groups were compared first in a Group (2) × Segment Type (2) repeated-measures ANOVA. Results revealed no main effect of group, $F < 1$, and no Group × Segment Type interaction, allowing the RD and RD+ groups to be collapsed into a combined RD group for subsequent analyses. A Group (2) × Segment Type (2) repeated-measures ANOVA, comparing the combined RD group with the CA controls, revealed a main effect of group, $F(1, 58) = 5.49$, $p < .05$, a main effect of segment type, $F(1, 58) = 17.73$, $p < .001$, but no interaction between these two factors. Thus, the combined RD group performed worse at first phoneme identification than did the CA control group (Table 3).

The combined RD group was compared with the RA control group in a Group (2) × Segment Type (2) repeated-measures ANOVA. Results showed a main effect of segment type, $F(1, 58) = 5.53$, $p < .05$, but no main effect of group and no Group × Segment

Table 3
Group means (and standard deviations) for lexical gating task variables

Group	120-ms segment[a]	Target word identification[b]
RD	12.8 (0.18)	93.6 (11.0)
RD+	12.9 (0.19)	95.3 (10.4)
CA controls	13.7 (0.25)	80.8 (6.8)
RA controls	12.3 (0.28)	87.5 (10.9)

Note. Standard deviations are in parentheses.

[a] Scores reflect the total number of first phonemes identified correctly out of 16.

[b] Scores reflect the percentages of the minimum gates required for correct word identification (16 words total).

Type interaction. Overall, these results suggest that the combined RD group is less able to make use of the acoustic information contained in the first segment as compared with CA controls and that its performance is indistinguishable from that of RA controls, who are approximately 4 years younger.

Because lexical gating tasks previously have focused on the number of gates required so as to correctly identify the target word, these data were also collected during the current study. The gate where the child correctly and consistently identified the target word was noted for each target word. Note that the participant needed to correctly identify the word at that gate and at subsequent gates for the lower numbered gate to count as the child's score. If the child guessed correctly at one gate but offered a different (incorrect) response at the following gate (which provided a longer segment of the word), the gate where the child initially guessed correctly did not count. This more conservative method was used because correct guesses followed by incorrect responses suggest that the child might not have fully interpreted the acoustic information provided by that particular segment so as to avoid a move away from that lexical entry when given a subsequent segment. The sum of the gate scores for the 16 target words was used in this secondary analysis. Results of the one-way ANOVA, comparing the combined RD, CA, and RA control groups, revealed a main effect of group, $F(2, 77) = 13.12$, $p < .001$. Pairwise comparisons (Tukey adjusted for multiple comparisons) indicated that the combined RD group required more gates to identify the word correctly than did both CA ($p < .001$) and RA ($p < .05$) controls.

Overall, the analyses of the lexical gating task support Hypothesis 1. That is, children with dyslexia, irrespective of their SSD or SLI history status, are less accurate in identifying the first phoneme of a target word when given a brief initial segment of that word than are their chronological age-matched peers. On this aspect of the task, their performance is at the level of the younger RA controls, and this does not allow us to rule out the alternative explanation that the deficit is due to differences in reading experience. However, there were differences relative to both CA and RA controls on the amount of acoustic information (number of gates) required to identify the word correctly. The latter measure might not be as pure an indicator of phonological representations because other higher order cognitive processes might be involved in accessing and retrieving the correct lexical entry compared with those that might be involved in arriving at the correct first phoneme.

Analysis of the Priming I task

Four related scores were obtained during the Priming I (congruent) and Priming II (incongruent) tasks: the number of correctly identified target words, the number of correctly identified first phonemes (regardless of word accuracy), the response time for correctly identified target words, and the response time for correctly identified first phonemes. All of these scores were obtained for each child in each priming condition (neutral tone, 120 ms, stretched, and 240 ms); the resulting group means for the Priming I experiment are shown in Table 4. Again, the stretched prime information is provided only because it was part of the experiment as administered; however, it is not interpreted here. We obtained phoneme accuracy data for two reasons. First, they give credit to those responses that had the correct first phoneme but an incorrect word. The latter situation would suggest that the prime activated the correct phonemic representation for the initial sound even though other phonological or lexical processes interfered with the selection of the rest of the target word. Second, using phoneme-level data in the priming experiment was analogous to the way the lexical gating task data were analyzed.

To ensure that priming was indeed taking place when the children heard a speech prime (vs. the nonspeech complex tone), a composite score composed of the average of the scores of the three speech prime types was computed. This composite was compared with the neutral prime score as a within-participants factor. A main effect of prime, $F(1, 76) = 80.90$, $p < .001$, was noted in a Group (4) × Condition (2) repeated-measures ANOVA with word accuracy as the dependent measure. No main effect of group or interaction was found. The exact same pattern of results was obtained with the other three dependent variables. Thus, the various speech primes were facilitating participants' access to the correct word or phoneme as compared with the neutral condition.

The RD and RD+ groups were compared next in a series of Group (2) × Prime (4) repeated-measures ANOVAs, one for each of the dependent variables. Results indicated that there were no main effects of group and no Group × Prime interactions in any of the analyses; therefore, the RD and RD+ groups were collapsed for all subsequent analyses.

Table 4
Group means (and standard deviations) for accuracy and latency variables by priming condition in Priming I task

Priming condition	RD	RD+	CA controls	RA controls
Word accuracy				
Neutral	12.4 (4.0)	12.8 (3.5)	12.0 (2.5)	11.2 (3.1)
120 ms	12.8 (2.8)	13.6 (2.8)	16.9 (2.6)	15.1 (1.8)
240 ms	18.1 (2.0)	17.9 (2.9)	16.9 (3.3)	16.7 (2.5)
Phoneme accuracy				
Neutral	15.9 (3.5)	17.0 (3.8)	17.3 (2.4)	15.7 (3.2)
120 ms	18.5 (2.6)	19.7 (2.9)	21.1 (1.8)	20.0 (1.9)
240 ms	22.8 (0.9)	22.5 (1.2)	22.3 (1.7)	22.2 (1.6)
Word RT				
Neutral	1231.3 (215.0)	1154.4 (155.3)	1137.8 (169.9)	1245.9 (148.9)
120 ms	1210.9 (195.4)	1114.9 (147.2)	1040.7 (106.4)	1176.7 (127.7)
240 ms	1122.1 (174.9)	1052.3 (114.9)	1024.6 (108.7)	1148.2 (112.1)
Phoneme RT				
Neutral	1313.8 (212.6)	1220.2 (164.1)	1233.3 (226.6)	1311.1 (176.9)
120 ms	1278.4 (206.2)	1187.1 (139.6)	1092.3 (133.4)	1227.7 (116.4)
240 ms	1172.2 (181.6)	1095.6 (145.6)	1079.1 (125.8)	1202.2 (131.8)

Note. Accuracy scores reflect the total number correct in each priming condition out of 24. RT, response time in milliseconds. Standard deviations are in parentheses.

Subsequent analyses for the Priming I experiment were performed on all four dependent variables; however, because word- and phoneme-level analyses showed identical patterns of results for accuracy as well as latency, only the phoneme-level data are reported here.

A series of 2 × 4 repeated-measures ANOVAs were conducted, comparing the combined RD group and the CA control group on the four dependent measures. For phoneme accuracy, a main effect of prime was obtained, $F(3, 134) = 77.98$, $p < .001$, as was a Group × Prime interaction, $F(3, 134) = 3.02$, $p < .05$. A trend for a main effect of group was also obtained, $F(1, 58) = 3.32$, $p = .07$. For phoneme RT, a main effect of prime, $F(2.6, 152) = 25.25$, $p < .001$, and a Group × Prime interaction, $F(2.6, 152) = 4.13$, $p < .01$, were obtained. (The Greenhouse–Geisser correction to the degrees of freedom for violation of the sphericity assumption was used in these particular analyses.)

The Group × Prime interactions were driven, in all cases, by the difference at the 120-ms prime condition, where the combined RD group, in contrast to the CA control group, did not show any improvement in performance relative to the neutral prime condition. This was corroborated by a series of *t* tests, performed for each dependent measure at each prime condition, comparing the combined RD and CA control groups. In each case, the only significant *t* test was at the 120-ms prime condition, $ts(58) \geqslant 2.83$, $p < .01$.

The same Group (2) × Prime (4) repeated-measures ANOVA analyses were conducted comparing the combined RD group with the RA control group. For phoneme accuracy, a main effect of prime was obtained, $F(3, 134) = 77.98$, $p < .001$, as was a trend for a Group × Prime interaction, $F(2.3, 132) = 2.52$, $p = .07$. There was no main effect of group. It should be noted that the same Group × Prime interaction was significant for word-level accuracy data, $F(2, 118) = 6.47$, $p < .01$. For phoneme RT, a main effect of prime, $F(2.4, 141) = 20.05$, $p < .001$, and a Group × Prime interaction, $F(2.4, 141) = 4.34$, $p < .01$, were obtained.

When the combined RD group was compared with the RA control group on phoneme accuracy and latency variables, the significant Group × Prime interactions obtained were again driven primarily by the results at the 120-ms prime condition. In the other three priming conditions (neutral, stretched, and 240 ms), the combined RD group outperformed the RA control group (i.e., the combined RD group has a higher accuracy mean, or a lower RT mean, than the RA control group). In contrast, the combined RD group underperformed relative to the RA control group at the 120-ms prime condition for both phoneme accuracy and latency. Although the difference in means did not reach significance with phoneme-level data, it did when the same comparison at the 120 ms prime condition was made using word-level data, $t(58) = 2.67$, $p < .01$.

Overall, the Priming I results support Hypothesis 1. The groups with dyslexia were not able to make use of the 120-ms prime as well as the CA or RA controls to facilitate speed of access or identification of the first phoneme or target word. The groups with dyslexia did, however, make use of the syllabic primes to facilitate first phoneme and target word identification.

Analysis of the Priming II task

A set of analyses similar to those performed with the Priming I data were conducted to analyze the incongruent priming task (Priming II). If the congruent primes were indeed causing the facilitatory effect on accuracy and latency scores in the Priming I task, mismatching the primes to target words should produce an interference effect on those same dependent measures. In this task, the expectation was for the CA and RA control groups to demonstrate a greater interference effect at the 120 ms prime condition relative to the RD groups. The analyses of the Priming II data revealed the expected pattern (Table 5); thus, the results are summarized briefly.

The RD and RD+ groups were compared on the four dependent measures via Group (2) × Prime (4) repeated-measures ANOVAs. There was no significant Group × Prime interaction on any of the four dependent measures; thus, the RD and RD+ groups were collapsed into one group (combined RD) for further analyses.

A series of Group (2) × Prime (4) repeated-measures ANOVAs were conducted comparing the combined RD group with the CA control group (and subsequently the RA control group) on the four dependent measures of interest. There were no significant Group × Prime interactions on either of the accuracy variables (word or phoneme) for any of the analyses. This may be due to the fact that by the time children were given the Priming II task, they had already heard the noise-masked target words twice before. Thus, practice effects may have masked or minimized any Group × Prime interactions. For word RT, however, there was a significant Group × Prime interaction, $F(3, 159) = 10.61$, $p < .001$, in the combined RD versus CA control comparison and a significant Group × Prime interaction, $F(2.6, 137) = 4.56$, $p < .05$, in the combined RD versus RA control comparison. The same pattern was observed for phoneme RT, with a significant Group × Prime interaction, $F(3, 136) = 5.21$, $p < .01$, in the CA comparison and a significant Group × Prime interaction, $F(2.5, 134) = 4.69$, $p < .01$, in the RA comparison.

For the phoneme-level RT data, the CA and RA control groups outperformed the combined RD group (i.e., had lower RTs) for each of the neutral, stretched, and 240-ms prime

Table 5
Group means (and standard deviations) for accuracy and latency variables by priming condition in Priming II task

Priming condition	RD	RD+	CA controls	RA controls
Word accuracy				
Neutral	6.7 (1.8)	6.6 (1.7)	7.6 (1.5)	6.8 (1.4)
120 ms	7.5 (1.8)	7.1 (1.8)	7.3 (1.9)	6.9 (2.2)
240 ms	6.4 (1.6)	6.3 (1.6)	5.7 (2.3)	5.9 (2.0)
Phoneme accuracy				
Neutral	8.5 (1.6)	9.3 (1.3)	9.6 (1.0)	8.8 (1.7)
120 ms	8.8 (1.6)	9.3 (1.3)	9.4 (1.5)	8.4 (1.6)
240 ms	8.2 (1.5)	8.3 (1.4)	7.9 (1.8)	8.0 (1.8)
Word RT				
Neutral	1135.3 (163.3)	1044.6 (112.8)	1011.7 (169.4)	1103.8 (168.1)
120 ms	1155.5 (203.1)	1034.4 (125.7)	1180.6 (279.3)	1215.9 (216.8)
240 ms	1306.5 (244.5)	1208.8 (171.1)	1132.1 (179.2)	1267.9 (220.0)
Phoneme RT				
Neutral	1206.1 (198.5)	1093.0 (133.0)	1045.3 (170.3)	1129.7 (169.5)
120 ms	1192.6 (202.2)	1086.7 (124.1)	1209.4 (278.1)	1253.2 (217.1)
240 ms	1319.0 (217.4)	1232.5 (164.6)	1221.9 (243.4)	1310.9 (214.0)

Note. Accuracy scores reflect the total number correct in each priming condition out of 12. RT, response time in milliseconds. Standard deviations are in parentheses.

conditions, but they underperformed the combined RD group (i.e., had higher RTs) at the 120-ms prime condition. Testing the within-participant difference between the neutral and 120-ms prime conditions, a series of 2 × 2 (Group × Prime) repeated-measures ANOVAs, comparing the combined RD group with the CA controls (and then RA controls) yielded significant Group × Prime interactions across the board, $Fs(1, 53) \geqslant 17.26$, $p < .001$. In all instances, the control groups increased their latency significantly from neutral to 120-ms prime conditions, whereas the combined RD group did not.

The results of the Priming II task are congruent with those of the Priming I task, but only for the RT data. The same pattern of priming was observed in both cases, with the RD groups being unable to benefit from the shorter prime as much as the controls. All participants benefited from priming in general, with latencies increasing (more interference detected) relative to the neutral prime. It is unclear why the accuracy data did not show the same pattern of results. It is possible that after hearing the stimuli three times, practice effects were able to mask any Group × Prime Type interactions. In addition, given the age of the children, it is less plausible for the mismatched prime to cause an outright error, which would require a complete breakdown of the inhibitory process. A slower reaction time is more plausible because the participants suppress the erroneous activation caused by the incongruent prime to retrieve the correct response. The latency data in the incongruent condition replicates the pattern observed in the congruent priming condition, strengthening the validity of the priming task results overall.

Summary of results from the experimental tasks

The three experimental tasks, taken together, provide strong support for Hypothesis 1, namely that there are less mature phonological representations in participants with dyslexia relative to both CA and RA controls. The comparison with a reading age-matched control group is a rather conservative one, suggesting that the immature phonological representations in dyslexic individuals are even less mature than those in individuals who are significantly younger in age and of equal ability on single-word reading (as a gross index of reading experience). Although the lexical gating task did not provide as strong a case of deficient implicit phonological representations relative to RA controls, there was no result, statistically significant or otherwise, that provided evidence to the contrary. The fact that three distinct methods measuring phonological representations in an implicit manner were used, and that all three converged on the same general pattern of results, strongly suggests that there are deficiencies beneath the level of PA in this population. The next section focuses on ruling out more general potential alternative explanations for the results found so far. Subsequently, we explore the relation among implicit phonological representations, speech perception abilities, and other core reading processes.

Assessing the effects of potential confounds

An attempt to match the groups on PIQ and parental education was made during ascertainment of the sample, and statistically group differences did not reach significance. However, there were trends for parental education, PIQ, and ADHD symptoms that potentially could act as confounds for the results obtained on the experimental tasks. A series of correlations were computed between parental education, PIQ, and the ADHD total score and all of the dependent measures generated by the three experimental tasks.

Results yielded only a few significant correlations. Phonemic accuracy on both priming tasks was significantly negatively correlated to both the ADHD composite score, $r = -.24$, $p < .05$, and PIQ, $r = -.28$, $p < .05$. In addition, the difference between the proportions of consonant and syllable errors on the syllable similarity task was significantly positively correlated with years of maternal education, $r = .25$, $p < .05$. No other variables from the experimental tasks were significantly correlated with potential covariates.

The analyses that involved the relevant correlated dependent variables were run again with the ADHD composite and maternal education variables as covariates. In no case did the results obtained in the initial analyses change due to the introduction of covariates into the model.

Relation between implicit phonological representations and speech perception

The speech perception data were analyzed to test Hypothesis 2, namely that children with dyslexia would show greater reliance on the dynamic portion of the acoustic signal and that this less mature pattern of speech perception would be related to the differences found on the measures of implicit phonological representations.

In the speech perception task, four slope variables (i.e., one each for *sa*, *sha*, *su*, and *shu*) and two boundary shift variables (one for *sa*-*sha* and one for *su*-*shu*) were derived from the probit analysis of the raw data. The slopes directly measured to what extent the child used the static portion of the acoustic signal to decide what syllable he or she heard. The boundary shift variables were measured by subtracting the static noise frequencies from the two curves (e.g., *sa* and *sha*) at the points where participants provided 50% /s/ responses. This difference is the boundary shift created by the different transition dynamics inherent in syllables that begin with different consonants (*s* vs. *sh*). The extent of the boundary shift provided a measure of the degree to which the child was using the consonant–vowel transition dynamics to decide what he or she heard. As normal children mature, they use the static portion of the signal more (steeper slopes) and use the transition dynamics less (decreased boundary shift) in making perceptual decisions.

In our data, all of the slope variables were positively intercorrelated, as were the boundary shift variables, whereas all of the slope variables were negatively correlated with boundary shift variables. The latter was expected because reliance on one type of cue generally reduces reliance on the other type of cue. Two standardized composites were computed—one for “slope” and one for “boundary shift”—by averaging standardized versions of each variable in each set and then restandardizing the resulting averaged score.

To see whether the speech perception task was sensitive to diagnosis, a series of repeated-measures ANOVAs was conducted, with type of cue (slope or boundary shift) as the within-participants factor. The first analysis compared the two RD groups. Results revealed no main effect of group and no Group × Cue interaction, so the RD groups were combined into one group. A Group (2) × Cue (2) repeated-measures ANOVA, comparing the combined RD group with CA controls on the two speech perception variables, was conducted. Results revealed no significant main effect of cue type and no main effect of group, but they revealed a significant Group × Cue interaction, $F(1, 47) = 6.98$, $p < .05$. As can be seen in Fig. 3, the CA controls exhibited a larger slope but a smaller boundary shift as compared with the combined RD group. Subsequent independent sample t tests corroborated that the difference between the groups was statistically significant for slope,

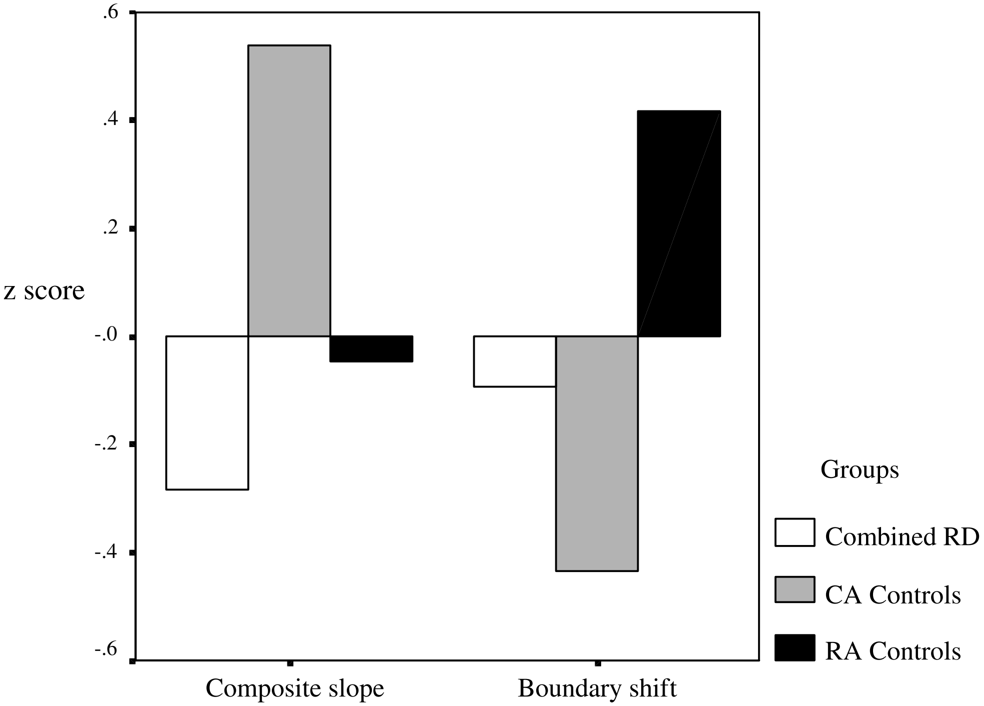

Fig. 3. Group means for composite slope and composite boundary shift on the speech perception task. Higher z scores on the composite slope and lower z scores on the boundary shift indicate a more mature speech perception pattern.

$t(53) = 2.91$, $p < .01$, but only a trend was observed for boundary shift, $t(53) = 1.88$, $p = .06$.

The same analyses were repeated comparing the combined RD group with the RA control group. The results showed a main effect of group, $F(1, 51) = 8.93$, $p < .01$, but no Group × Cue interaction. As before, the RA control group had a greater slope than the group with dyslexia, but it also displayed a greater boundary shift. This pattern suggests that the RA controls were relying more on both the static and dynamic portions of the speech signal than were the participants with dyslexia. Overall, as Hypothesis 2 predicted, participants with dyslexia showed relatively greater reliance on the dynamic portion of the speech signal and relatively less reliance on the static portion of the speech signal.

Prior to comparing the experimental tasks with the speech perception task, a composite implicit phonological representation score was computed. A factor score was derived by using one variable from each of the three experimental tasks. The proportion difference score (consonant–syllabic errors) from the syllable similarity task, the phoneme accuracy score at the 120-ms prime condition from the Priming I task, and the number of correct first phoneme identifications from the lexical gating task were subjected to a principal components factor analysis with an oblique (direct oblimin) rotation. The factor analysis solution was composed of one factor explaining 56% of the variance. The factor loadings for each variable were as follows: proportion difference score, .84; phoneme word accuracy, .85; and correct phoneme identification at the first segment, .52. The standardized factor score for each participant was used in subsequent analyses as an implicit phonological representation composite score. The relation between the speech perception variables and the composite implicit factor score was explored via a correlational analysis. As predicted in Hypothesis 2, higher performance on the composite implicit phonological representation factor score was related to a higher score on the slope measure, $r = .37$, $p < .01$. Only a trend was observed for the relation with the boundary shift measure, $r = -.27$, $p = .06$, although it was in the predicted direction (higher implicit composite score related to a lower boundary shift).

Relation between implicit phonological representations and PA, nonword repetition, and RAN

A set of analyses analogous to those conducted with speech perception were performed to investigate the relation between implicit phonological representations and the other core deficits in dyslexia (PA, nonword repetition, and RAN). For PA, a composite variable was computed by taking the standardized average of individual *z* scores on the Pig Latin and phoneme reversal tasks. For nonword repetition, the standardized total number of correct phonemes produced at the four-syllable word level was used as the dependent variable. This was the most discriminating of the nonword repetition variables. For RAN, a residualized composite of rapid color and object naming was used as the dependent variable with the effects of age removed.

In general, as shown in Table 6, groups with dyslexia performed worse than controls on all three types of measures. For PA, both RD groups performed significantly worse than CA controls, but only the RD+ group performed significantly worse than RA controls. For nonword repetition, the RD+ group performed significantly worse than both CA and RA controls, but the RD group did not differ from either control group, although the pattern of scores was in the predicted direction (the RD group was next worse after the RD+ group). Finally, for RAN, the RD+ group was significantly worse than CA controls. No other group differences were found on RAN, although the predicted pattern of scores was obtained once again.

A correlational analysis was conducted to investigate the relations among these variables dimensionally. A complete correlational matrix is shown in Table 7. Notably, the

Table 6
Group differences on PA, nonword repetition, and RAN

Dependent variable	RD group	RD+ group	CA controls	RA controls	*F* value
PA composite	$-0.04\ (0.72)_a$	$-0.91\ (0.98)_b$	$0.97\ (0.55)_c$	$-0.06\ (0.72)_a$	20.0^{***}
Nonword repetition	$0.12\ (0.66)_a$	$-0.79\ (1.25)_b$	$0.36\ (0.79)_a$	$0.27\ (0.84)_a$	6.60^{***}
RAN composite	$-0.02\ (0.81)_a$	$0.36\ (1.09)_b$	$-0.49\ (0.77)_a$	$0.14\ (1.11)_{ab}$	2.76^{*}
Articulation rate	$0.02\ (1.05)_a$	$-0.40\ (1.20)_b$	$0.39\ (0.79)_a$	$-0.03\ (0.81)_a$	$2.16^{\dagger}$

Note. All values are *z* scores with standard deviations in parentheses. Groups with no shared subscripts are significantly different at $p < .05$ or less. RAN and articulation rate composite scores are standardized residuals after removing the effects of age. RAN score is a latency measure, with more negative scores reflecting better performance.

[*] $p < .05$.

[***] $p < .001$.

[†] $p < .10$.

Table 7
Correlation matrix of implicit phonological representations and explicit reading process and outcome variables

	1	2	3	4	5	6
1. Implicit representations[a]	—					
2. Phoneme awareness	$.49^{**}$	—				
3. Composite slope	$.36^{**}$	$.29^{*}$	—			
4. Boundary shift composite	$-.23^{\dagger}$	$-.20$	$-.47^{**}$	—		
5. Nonword repetition	$.35^{**}$	$.43^{**}$	$.31^{**}$	$-.07$	—	
6. Rapid automatized naming	$-.15$	$-.30^{**}$	$-.09$	$-.01$	$-.29^{**}$	—
7. Reading composite	$.53^{**}$	$.55^{**}$	$.30^{*}$	$-.04$	$.48^{**}$	$-.32^{**}$

Note. All significance tests are two-tailed.

[a] Composite measure of the three experimental tasks; slope and boundary shift composites are from the speech perception task.

[*] $p < .05$.

[**] $p < .01$.

[†] $p < .10$.

composite implicit factor score was significantly correlated with PA, $r = .49$, $p < .001$, and nonword repetition, $r = .35$, $p < .01$, but not with RAN, $r = -.15$, $p = .22$. Thus, a more mature pattern of implicit phonological representations was related to higher phonological awareness ability and better short-term phonological memory but was not associated with speed of lexical retrieval. Finally, the implicit phonological representation factor score was significantly correlated with the reading composite, $r = .53$, $p < .001$.

Next, the implicit factor score was paired with each of the other proposed core deficits in dyslexia in a series of hierarchical regression analyses each time predicting the reading composite score. As can be seen in Table 8, all of the variables except the speech perception slope composite predicted unique variance in reading when entered in the second step of the regression model. By subtracting the appropriate semipartial correlations, we

Table 8
Summary of results of hierarchical regression analyses predicting the composite reading score

Regression analysis	Independent variables in each regression equation	Total R^2	R^2 change	F change	Percentage of total explanatory variance of IV$_1$ shared by implicit factor (IV$_2$)
1	IV$_1$ Speech perception	.38	.02	1.59	83
	IV$_2$ Implicit factor		.26	20.12^{***}	
2	IV$_1$ PA composite	.39	.11	12.80^{**}	63
	IV$_2$ Implicit factor		.09	9.89	
3	IV$_1$ Nonword repetition	.38	.10	11.30^{**}	57
	IV$_2$ Implicit factor		.15	16.70^{***}	
4	IV$_1$ RAN composite	.34	.06	6.20^{*}	41
	IV$_2$ Implicit factor		.24	25.20^{***}	

Note. R^2 change reflects the amount of unique variance contributed by that IV when entered *second* into the hierarchical regression equation. The last column answers the following question: “Of the total amount of variance that IV$_1$ accounts for in the reading composite, how much of it is shared with the implicit factor score?”

[*] $p < .05$.

[**] $p < .01$.

[***] $p < .001$.

computed that 83% of the explanatory variance of speech perception in predicting reading was shared by the implicit factor score. Although the implicit factor score shared a significant proportion of the explanatory variance of each of the other variables (PA, nonword repetition, and RAN), each of these also contributed unique variance in predicting reading. One may infer from these results that speech perception and implicit phonological representations are very related constructs.

In summary, implicit phonological representations shared a significant amount of variance with the reading composite score. All of the correlations between implicit phonological representations and other reading process variables that have been proposed as potential core deficits in dyslexia (PA, speech perception, and short-term verbal memory) were significant, with the notable exception of a lack of relation between implicit phonological representations and RAN.

Discussion

This study was designed to measure implicit phonological representations in children with dyslexia and relate this construct to various other levels of linguistic processing and literacy variables known to be affected in this disorder. Specifically, this study tested the segmentation hypothesis, which states that one of the core deficits in dyslexia, PA, is due to phonological representations that are less segmented than in children who read normally. This hypothesis has been difficult to test directly because most clinical and research tasks require a certain level of metalinguistic ability (i.e., actual awareness of phonemic structure or manipulation of phonemic segments). This study used three experimental tasks (a lexical gating task, a priming task, and a syllable similarity task) to measure the construct of implicit phonological representations. The surface requirements and characteristics of each task were different enough as to provide protection against the possibility that method variance could account for the converging results that were found.

Specifically, we found the predicted pattern of results on all three experimental tasks; children with dyslexia performed consistently worse than CA controls and generally worse than RA controls. These results provide strong support for less mature phonological representations in children with dyslexia as predicted by Hypothesis 1. The comparison with a reading age-matched control group is a rather conservative one, suggesting that the immature phonological representations in dyslexic individuals are even less mature than those in individuals who are significantly younger in age and of equal ability on single-word reading (as a gross index of reading experience). A factor analysis showed that the three most relevant dependent variables, one from each task, all loaded significantly on a latent construct of implicit phonological representations.

The participants with dyslexia in this study were also differentiated from CA and RA controls in the speech perception identification task. Both RD groups were less sensitive to the static portion of the synthesized stimuli than their chronological age- and reading age-matched peers. In addition, both RD groups relied on the dynamic portion of the speech stimuli to a greater degree than did CA controls but not RA controls. Hypothesis 2 was also supported by the significant correlation between the composite speech perception variables and the implicit phonological representation factor score. Children whose implicit representations were more segmental showed a more mature pattern of weighting cues in the speech perception labeling task.

Hypothesis 3 was also partly supported by the current set of results. The implicit phonological representations factor score was significantly correlated with two of the three other potential core deficits in dyslexia (PA and short-term verbal memory), and it shared a significant amount of variance with each of these variables in predicting a reading composite score. Implicit phonological representations were not significantly correlated with RAN, but the unique contribution of each to the prediction of the reading composite score remained statistically significant. The latter finding is consistent with recent results obtained by Pennington, Cardoso-Martins, Green, and Lefly (2001) that showed that rapid serial naming contributed independent variation to various literacy skills, but its contribution was more modest relative to the contribution of PA. Although PA shares approximately 9% of variance with RAN in the current study ($r = -.30$), the latter shares only a quarter of that amount of variance (2.5%) with implicit phonological representations. This suggests that although PA and RAN may have some general cognitive factors in common, they do not seem to share an underlying phonological deficit. This is consistent with findings from Pennington and colleagues' findings that RAN significantly predicts reading fluency but does less well in predicting phonological coding ability.

Hypothesis 4 was generally not supported by the results obtained in this study because there were very few differences between the RD and RD+ groups across most tasks. There are a few possible reasons for this outcome. First, the basis on which participants were chosen to be in the RD+ group was a retrospective questionnaire filled out by their parents, asking relevant questions about speech and language symptoms either that were currently manifested by the children or that were present when the children were younger. Because most articulation disorder symptoms normalize by 6 years of age, the accuracy and reliability of the retrospective data may be less than optimal given that the average age of the participants with dyslexia in this study was slightly more than 12 years. Not only may the categorical distinction between an articulation disorder and normal speech be less accurate when using retrospective data supplied by parents, but also the severity of the articulation deficit is difficult to judge. Some parents may interpret a mild distortion as an articulation deficit, whereas others may require a much higher level of speech unintelligibility to ascribe the presence of a disorder.

Parents were also asked about their children's language deficits. Although symptoms of SLI can persist in children through adolescence (Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998), the deficits are generally less apparent to parents, who may judge their children's linguistic skills on general conversational abilities rather than on more complex linguistic tasks. Thus, there is an inherent amount of unreliability in this type of questionnaire measure that may have affected the group selection process, resulting in a less severe comorbid group of children with dyslexia.

Another factor that may have decreased the ability to find differences between the RD and RD+ groups is the fact they were matched on severity of reading disability. There is some recent evidence that suggests that individuals with dyslexia who are comorbid for SSD/SLI often are more severely affected in reading than are those without a comorbid diagnosis (Harm & Seidenberg, 1999; Tunick & Pennington, 2002). By matching on severity of reading disability, the current RD+ group could have been selected from the top end of the RD+ population distribution, yielding a less severe sample of RD+ individuals.

For the reasons just listed, it is possible that the quasiexperimental manipulation of SSD/SLI symptoms was not as successful as we had hoped it would be. It remains possible that a more severe group of comorbid dyslexic individuals would have performed worse than the RD-only group on the experimental tasks. Consistent with this possibility, exploratory analyses using the total score on the speech and language questionnaire were conducted to assess whether using a continuous measure of SSD/SLI rather than a categorical one would reveal the expected relations. The results showed that the continuous measure of speech and language impairment was significantly correlated with the implicit phonological representation composite measure. In addition, the continuous measure of speech and language impairment accounted for a significant amount of unique variance (controlling for implicit phonological representations) in predicting both PA and the reading composite score.

Alternative explanations for deficient phonological representations

Differences in IQ, attentional capacities, lexical–semantic ability, and demographic characteristics (e.g., parental education and SES) were considered as potential alternative explanations for the current results. Matching was conducted in an attempt to make the groups equivalent on some of these relevant variables. The groups were selected to be similar on age, IQ, gender, parental education and SES, and severity of reading deficits. Although there were no significant statistical differences on any of the variables, trends did exist for parental education and PIQ. Furthermore, there was a trend for group differences in ADHD symptoms, a variable that was not part of the matching process. All of the trends were in the expected direction, with the control groups being characterized by higher PIQ and parental education and lower levels of ADHD symptoms.

Only a few correlations were significant between demographic and dependent measures. One of these was unexpected; PIQ was negatively correlated with letter accuracy in the 120-ms prime condition of the Priming I task in all four groups. Given the direction of effect, however, this served only to accentuate group differences when PIQ was used as a covariate. Notwithstanding the expected correlations between parental education and ADHD symptoms and experimental dependent variables, the results of the analyses of covariance did not change the pattern of the original results. Thus, even when the effects of these potential confounds were taken into account, the deficit in implicit phonological representations in participants with dyslexia remained.

A more direct potential confound stems from one similarity across all three experimental tasks, namely that they all require children to store or access information from the lexical system. To the extent that participants with dyslexia have a less robust set of connections that make up the lexical–semantic network, or to the extent that they have different semantic knowledge bases, one might posit that the results of the experimental tasks reflect those differences rather than deficits at the phonological level.

Although this criticism cannot be dismissed entirely, there are a few reasons to suspect that the effects of such a confound are minimal. First, all of the words selected for the lexical and priming tasks were high-frequency words with age-of-acquisition levels between 2 and 3 years of age. Thus, all participants would be expected to be highly familiar with the stimuli and thus to have a robust lexical representation for them. Second, the words in the priming task were paired with different target words to create four distinct versions of the task. The versions were counterbalanced within groups. Therefore, only a global deficit in lexical retrieval would be able to explain the same pattern of results across all four versions. If that were the case, one would expect main effects of group rather than Group × Prime Type interactions. Finally, only the RD+ group showed a difference in

RAN, which provides a measure of lexical retrieval efficiency. However, when a measure of articulatory speed was taken into account, the difference between the RD+ and CA control groups on the RAN composite variable no longer was significant. Thus, the RD groups are not manifesting lexical retrieval difficulties for colors and objects, the two series of RAN stimuli that are learned at an early age.

Another potential alternative explanation, related to the lexical retrieval issue, is the fact that there are group differences in articulatory rate (i.e., the RD+ group has a slower articulation rate) and that the latter is significantly correlated with the phonological representations factor score. Because both the lexical gating and priming tasks required a verbal response (as opposed to the syllable similarity task, which required only pointing), differences in accessing and regulating articulatory output may have interacted with phonological representations in such a manner as to create spurious differences in the priming and lexical gating tasks. To rule out this potential confound, articulatory rate was used as a covariate in a series of repeated-measures analyses of covariance (ANCOVAs) performed for the main experimental analyses (excluding the syllable similarity task). The pattern of results for each experimental task did not differ from that obtained originally. The fact that the syllable similarity measure was significantly correlated with the relevant variables from lexical gating and priming tasks, even though no articulatory output was required, also bolsters the argument that the deficits in phonological representations cannot be explained by differences in articulatory rate.

There is the possibility that there were differences in strategic approach used by participants with dyslexia and controls when doing the experimental tasks. Like the potential lexical confound discussed earlier, this also is difficult to rule out completely. However, the multiple methods used to obtain convergence on the implicit phonological representation construct mitigates the potential threat of this confound. The surface characteristics of each experimental task are quite different, and it is unlikely that similar strategies would be used across all three tasks. It would be even less likely for participants in the groups with dyslexia to somehow use similar strategies among themselves so as to cause the group differences observed across all three tasks.

Two other potential criticisms regarding sample characteristics (twining and reading severity level) and their potential effects on the generalizability of the results require specific mention. First, because the RD participants in this study were drawn from twin pairs, one could claim that dyslexic twins are not representative of the larger population of singleton children with reading difficulties. The impact of twining typically is quite subtle, however, and both the current sample and the larger twin sample from which it was derived manifest all of the core dyslexic deficits seen in singleton participants with dyslexia. In addition, the demographic characteristics normally associated with dyslexia, including lower parental education, lower PIQ, and increased symptom levels of ADHD, are also present. Thus, we believe that our current sample does not differ significantly from the general (singleton) population of individuals with dyslexia.

Second, the overall severity of the RD sample in this study is mild to moderate. Although there is an ample range of reading composite scores, the mean reading composite score for the current sample is only 1 *SD* below the discrimination function cutoff of zero. Nevertheless, the mean composite reading score is nearly identical to the overall mean in the dyslexic sample of twins in the CLDRC study. The reading composite means of the RD and RD+ groups were −1.01 ($SD = 0.65$) and −1.18 ($SD = 0.81$), respectively, whereas the dyslexic proband reading composite mean in the entire twin study was −1.08

($SD = 0.81$). This suggests that the current sample is at least representative of the larger twin sample with regard to severity. It should also be noted that reading scores were obtained 1 to 3 years prior to the current study; therefore, the RD participants' reading skills most likely had improved (via practice, maturation, and/or intervention) since that time. The RD group still showed a significant discrepancy from CA controls in single-word reading, however, at the time of the study. It is informative, then, that despite the fact that most of these participants with dyslexia had been receiving intervention, there were still large differences observed in implicit phonological representations relative to CA and RA controls.

Relation of current findings to previous studies of implicit phonological processing

The current findings support previous literature that has suggested that individuals with reading disability have poor phonological representations and adds to the literature in a number of ways. As in Matsala (1997), this study also found that reading-disabled children required more of the target word to identify it relative to CA controls on the lexical gating task. The current study also found this difference compared with RA controls. Because the current cohort of participants with dyslexia was 3 years older than those studied by Matsala, the deficits appear to be persistent into early adolescence. Departing from Matsala's procedures, this study also demonstrated that children with dyslexia have more difficulty in identifying the correct first phoneme at the earliest gated segment (relative to CA controls), suggesting that the deficit is not only one of lexical access and word recognition but also one of phoneme recognition. The current results are also quite similar to the findings of Dollaghan (1998), who administered the lexical gating task to children with SLI and found that first phoneme recognition was also deficient. The consistency with Dollaghan's results also suggests an overlap in phonological representation deficits between individuals with dyslexia and individuals with SLI. This would not be surprising given the comorbidity that has been documented between these two disorders (Aram, Ekelman, & Nation, 1984; Gallagher et al., 2000; Hall & Tomblin, 1978; Rutter & Mahwood, 1991) and the possibility that they may share, in part, a common etiology. The latter supposition currently is under investigation, but preliminary analyses of twin data at the CLDRC (Hulslander & Olson, 1999) showed bivariate heritability between reading and nonword repetition, with the latter being a phenotypic marker for SLI that is under genetic influence.

The results of the current study are also consistent with those of Lance et al. (1997), who found that children with reading disability had poorer "implicit phonological awareness." The current study extends Lance and colleagues' findings by using tasks where the child could not resort to a phonetic comparison of the stimulus to Standard English phonetic rules. The more holistic level of processing that could account for Lance and colleagues' findings cannot be invoked as an alternative explanation of the results in the current study.

Our results are not consistent with those of Griffiths and Snowling (2001), who also measured word identification skills in children with dyslexia using a similar lexical gating procedure. Griffiths and Snowling did not find differences between participants with dyslexia and controls (CA or RA) for the amount of a word required for correct identification of the target stimuli despite differences found between groups on nonword reading and rapid naming skills. There are a number of differences between the two studies that warrant discussion. First, the ascertainment criteria are somewhat dissimilar, with the current study using a discrimination function that incorporates single-word reading, spelling, and reading comprehension rather than just single-word reading. One wonders whether the discrimination function identifies participants with a slightly more impaired or broadly defined profile than those selected using only a single-word reading recognition measure. Second, no lower limit was imposed on vocabulary level in our study (because we were recruiting children with comorbid speech/language impairment), whereas Griffiths and Snowling used a lower cutoff of 8 on the WISC-III Vocabulary subtest and matched their groups on this variable. By selecting participants with slightly higher than average vocabulary skills, Griffiths and Snowling may have decreased their chances of finding differences on the lexical gating task if one follows the reasoning that increased lexical growth affects segmentation of phonological representations. Third, by looking only at time to word identification, Griffiths and Snowling were not able to ascertain whether differences existed in identifying the first phoneme after presentation of the first gate. This, we argue, is a more rigorous test of the segmentation hypothesis because the average duration (350–400 ms) of the acoustic information required to identify the entire word corresponds to at least the first consonant–vowel segment, which is much closer to a syllabic level of segmentation. Individuals with dyslexia might not have difficulty in using acoustic information to access syllabic or word-level information, thereby reducing the chances of finding differences when time to word identification is used.

It should be noted that when our data were analyzed using the number of gates required to identify the word, we still found differences between participants with dyslexia and CA and RA controls. These results may be due to lower vocabulary skills in our children, the additional speech and language deficits for the RD+ group, and the fact that we used only target words with sparse neighborhood densities to maximize the possibility of finding differences. When one looks at the Griffiths and Snowling (2001) data on high-frequency/sparse-density words only, however, there are still no observable group differences.

Griffiths and Snowling (2001) also used a "duration blocked" method to present stimuli (to prevent response perseveration and negative feedback), whereas the current study used a "stimulus blocked" format of presentation. Although the argument for using a duration blocked format is sound, the fact that our lexical gating results related in the expected manner with the other implicit and explicit phonological tasks in our study suggests that our results, obtained via the stimulus blocking method, were not likely spurious. The scoring method for determining the amount of acoustic phonetic input required to identify a word was similar in the two studies because both required the word to be correct on subsequent gated intervals. Because all of our participants, including the RA controls, were able to identify the word by the last gate, the issue of disqualifying stimuli or adding 50 ms if a word was not identified at all was not an issue. When we transform our data from gates to milliseconds, our range of scores is similar to Griffiths and Snowling's data when comparing directly with their sparse-frequency/high-frequency subgroup of words (e.g., 361.8 ms in our study vs. 384 .0 ms in Griffiths and Snowling's study for CA controls). Our slightly lower time is likely due to the difference in the average length of the words used (Ms = 560 vs. 634 ms).

Finally, our study corroborates findings by de Gelder and Vroomen (1991), who found differences between adults with dyslexia and normal controls on a task similar to the syllable similarity task used in this study. The current results indicate that even with the added variability that might be expected when testing children, the results of the syllable similarity task are still indicative of an immature pattern of encoding.

Relation of current findings to previous studies of speech perception

The developmental weighting shift hypothesis, proposed by Nittrouer (1992), captures a developmental phenomenon in speech perception abilities by demonstrating that there is maturation involved in what acoustic cues individuals will rely on when performing categorical perception experiments. In addition, Nittrouer showed that this maturation is related to measures of literacy, articulation ability, and PA. The link that appears to be implied by such relations is that the development of more mature patterns of speech perception is most likely (bidirectionally) related to the development of more mature representational patterns for the phonemes and words that are the end result of analysis of the speech stream. To create a psychological reality (phonemes and words) from a physical stimulus (sound waves propagating in air), the brain needs to find an efficient method of abstracting the invariant characteristics of the speech signal. The current study was able to show more explicitly that a strong association between maturation of perceptual abilities and phonological representations indeed exists.

Our current results are also somewhat consistent with those of Manis et al. (1997) in that not all of the participants with dyslexia showed abnormal speech perception scores. In fact, 7 of 35 participants had slope scores that were above the CA mean. These 7 children with dyslexia, however, were still significantly worse than CA controls at PA, implicit phonological representations, and reading, suggesting the possibility of another route to poor phonological representations other than poor speech perception. Analyzing the sample in the same way as Manis and colleagues, in our study there were 12 of 40 children with dyslexia who demonstrated normal PA (defined as scores not less than −1.0 *SD* from the CA mean). These 12 children were still impaired on implicit phonological representations and reading measures compared with CA controls. Although not significantly impaired on speech perception variables as a group, 40% of those with normal PA skills still scored in the abnormal range on the slope score for the speech perception task. These results suggest that some children with dyslexia can remediate their PA skills but that their implicit phonological representations and speech perception skills continue to be a residual deficit.

Finally, it should be noted that the current study does not speak directly as to what caused the underlying implicit phonological representation deficit. We do suggest that a relation exists between one measure of categorical perception and implicit phonological representation, but this study did not specifically address whether either one or both of these deficits may be partly attributable to the effects of a more general deficit in auditory processing.

Implications for future research

A number of questions raised by the current study could be the impetus for further research. First, replication of the current study with younger dyslexic children could help to determine whether the phonological representation deficit is present during the early stages of reading development and how much of it is affected by reading experience.

Because PA is known to have a reciprocal relation with reading, it is hypothesized that implicit phonological representations may show the same type of relation. Pursuing research with adults with dyslexia, who may have compensated for their reading deficits to a larger degree, would be fruitful as a way of measuring the persistence of such a deficit and its relation to other residual difficulties (e.g., spelling ability).

Another line of inquiry would be to investigate the heritability of underlying phonological representations. It is hypothesized that this trait, like other complex cognitive phenotypes, would be subject to both genetic and environmental influences. Understanding these influences and how they might relate to the genetic influences on other reading and reading-related variables may help to form a more cohesive picture of how dyslexia develops and interacts with other contextual factors. Finding that these measures of implicit phonological representations are coheritable with PA and reading measures would also provide incremental validity for the current construct.

Pursuing the possibilities in the genetic realm even further, a heritable underlying deficit strongly related to reading could be used as an even more refined phenotypic marker for molecular genetic studies attempting to identify the genes that influence reading and dyslexia. Quantitative trait loci (QTL) have already been identified, in replicated studies, on a number of chromosomes (Cardon et al., 1994; Fisher et al., 1999; Gayan et al., 1999; Grigorenko et al., 1997). This type of phenotypic marker could aid the identification of actual genes in the QTL regions identified so far.

Currently, there are also various studies investigating the genetic and phenotypic overlap between dyslexia and speech and language disorders. The efforts to find common core deficits may benefit from more discrete and specific phenotypes. Depending on how well the construct of implicit phonological representations holds up in replication studies, it could potentially serve this purpose.

At yet another level of analysis, functional neuroimaging studies involving the implicit phonological representation construct may provide a clearer picture of brain activation differences in good and poor readers. This type of investigation could be extended across languages because the deficit in implicit phonological representations may remain even when the transparency and grain size of the orthography minimizes the functional impact of the disorder (Landerl, Wimmer, & Frith, 1997).

Finally, it would be important to measure the effect of current remediation paradigms on underlying phonological representations, so that we can better understand which targets are benefiting from different kinds of intervention. It may well be that changes in implicit phonological representations are a good marker of change in intervention. In contrast, reading and PA skills may improve without much change in the segmentation or robustness of these implicit representations. As discussed previously, there are instances where neither PA deficits nor speech perception deficits are apparent in dyslexic individuals, yet a measurable discrepancy in the underlying phonological representation remains. It is hoped that understanding which aspects of cognitive and linguistic processing are necessary and sufficient for improved functional outcomes will help reading remediation specialists and educators to focus on those aspects that are the most conducive to reading progress. To achieve this level of understanding, however, the relevant constructs at different levels of phonological and cognitive processing need to be measured validly and accurately. The current tasks, or modifications thereof, may provide a basis for this type of measurement at the implicit phonological representation level.

Appendix

Syllable similarity task stimuli

Set 1:	/bIs/	/diz/	/bun/
Set 2:	/pIm/	/pus/	/ten/
Set 3:	/zaid/	/seb/	/sʊk/
Set 4:	/tʌp/	/pech/	/tash/
Set 5:	/faem/	/fup/	/sʊt/

Lexical gating task stimuli

(practice words: Doll, Rock, Nail)

Boat
Bead
Bus
Dish
Dirt
Dog
Girl
Game
Moose
Mouth
Watch
Wood
Feet
Fork
Light
Rain

Priming task stimuli

(practice words: Gate, Ball, Lock, Meat, Den, Bean, Chair, Sun)

Bar
Beach
Book
Bike
Bird
Belt
Deer
Door
Dust
Dime
Dart
Deck
Gold
Guess
Gun
Gas
Gun
Goat
Church
Voice
Wind
Wood
Soap
Farm
Lamp
Fish
Shirt
Cheese
Note
Juice
Food
Soup
Milk
Mud
Room
Jar
Face
Knife
Moon
Yard
Rope
Worm
Bush
Bath
Duck
Desk
Gut
Geese

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Aram, D. M., Ekelman, B. L., & Nation, J. E. (1984). Preschoolers with language disorders: 10 years later. *Journal of Speech and Hearing Research, 27*, 232–244.

Bird, J., & Bishop, D. V. M. (1992). Perception and awareness of phonemes in phonologically impaired children. *European Journal of Disorders of Communication, 27*, 289–311.

Bird, J., Bishop, D. V. M., & Freeman, N. H. (1995). Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research, 38*, 446–462.

Brady, S., Shankweiler, D., & Mann, V. A. (1983). Speech perception and memory and coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*, 345–367.

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

de Gelder, B., & Vroomen, J. (1991). Phonological deficits: Beneath the surface of reading acquisition problems. *Psychological Research, 53*, 88–97.

DeFries, J. C. (1985). Colorado Reading Project. In D. Gray & J. Kavanaugh (Eds.), *Biobehavioral measures of dyslexia* (pp. 107–122). Parkton, MD: York Press.

DeFries, J. C., Filipek, P. A., Fulker, D. W., Olson, R. K., Pennington, B. F., Smith, S. D., et al. (1997). Colorado learning disabilities research center. *Learning Disability Quarterly, 8*, 7–19.

Denckla, M. B., & Rudel, R. G. (1976). Rapid automatized naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471–479.

DeWeirdt, W. (1988). Speech perception and frequency discrimination in good and poor readers. *Applied Psycholinguistics, 9*, 163–183.

Dollaghan, C. A. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics, 19*, 193–207.

Dollaghan, C. A., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136–1146.

Fisher, S. E., Marlow, A. J., Lamb, J. M. E., Williams, D. F., Richardson, A. J., Weeks, D. E., et al. (1999). A quantitative-trait locus on chromosome 6p influences different aspects of developmental dyslexia. *American Journal of Human Genetics, 64*, 146–156.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 97–117). Hillsdale, NJ: Lawrence Erlbaum.

Gallagher, A. M., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41*, 202–213.

Gathercole, S. E., & Baddeley, A. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439–454.

Gayan, J., Smith, S. D., Cherny, S. S., Cardon, L. R., Fulker, D. W., Brower, A. M., et al. (1999). Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *American Journal of Human Genetics, 64*, 157–164.

Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology, 32*, 401–424.

Griffiths, Y. M., & Snowling, M. J. (2001). Auditory word identification and phonological skills in dyslexic and average readers. *Applied Psycholinguistics, 22*, 419–439.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., et al. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *American Journal of Human Genetics, 60*, 27–39.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*, 267–283.

Hall, P. K., & Tomblin, B. (1978). A follow-up study of children with articulation and language disorders. *Journal of Speech and Hearing Disorders, 43*, 227–241.

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

Hulslander, J., & Olson, R. (1999). *Nonword repetition: Phenotypic and genetic relations in reading disabled children.* Poster presented at the meeting of the Society of the Scientific Study of Reading, Montreal, Canada.

Katz, R. (1986). Phonological deficiencies in children with reading disabilities: Evidence from an object naming task. *Cognition, 22*, 225–257.

Kucera, H., & Francis, N. (1967). *Computational analysis of present day American English*. Providence, RI: Brown University Press.

Lance, D. M., Swanson, L. A., & Peterson, H. A. (1997). A validity study of an implicit phonological awareness paradigm. *Journal of Speech and Hearing Research, 40*, 1002–1010.

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German–English comparison. *Cognition, 63*, 315–334.

Lewis, B., & Freebairn, L. (1992). Residual effects of preschool phonology disorders in grade school, adolescence, and adulthood. *Journal of Speech and Hearing Research, 35*, 819–831.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., et al. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology, 66*, 211–235.

Markey, K. (1994). *The sensorimotor foundations of phonology: A computational model of early childhood articulatory and phonetic development.* Ph.D. Dissertation, University of Colorado.

Matsala, J. L. (1997). Spoken word recognition in reading disabled children. *Journal of Educational Psychology, 89*, 159–169.

Messer, S. (1967). Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior, 6*, 609–613.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *Journal of Phonetics, 20*, 1–32.

Nittrouer, S. (1996). The relation between speech perception and phonemic awareness: Evidence from low-SES children and children with chronic OM. *Journal of Speech and Hearing Research, 39*, 1059–1070.

Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems? *Journal of Speech, Language, and Hearing Research, 42*, 925–942.

Nittrouer, S., & Studdert-Kennedy, K. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research, 30*, 319–329.

Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 1–64). San Diego: Academic Press.

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

Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Development, 61*, 1753–1778.

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72*, 816–833.

Pennington, B. F., Cardoso-Martins, C., Green, P., & Lefly, D. (2001). Comparing the phonological and double deficit hypothesis for developmental dyslexia. *Reading and Writing, 14*, 707–755.

Reed, M. A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology, 48*, 270–292.

Rescorla, L. (1989). The Language Development Survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders, 54*, 587–599.

Reznick, J., & Goldsmith, L. (1989). A multiple form word production checklist for assessing early language. *Journal of Child Language, 16*, 91–100.

Rutter, M., & Mahwood, L. (1991). The long-term psychosocial sequelae of specific developmental disorders of speech and language. In M. Rutter & P. Casaer (Eds.), *Biological risk factors for psychosocial disorders* (pp. 233–259). Cambridge, UK: Cambridge University Press.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728–1743.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In A. C. P. Accardo & B. Shapiro (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75–119). Timonium, MD: York.

Serniclaes, W., Van Heghe, S., Mousty, P., Carre, R., & Sprenger-Charolle, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology, 87*, 336–361.

Snowling, M. (1981). Phonemic deficits in developmental dyslexia. *Psychological Review, 43*, 219–234.

Snowling, M., Goulandris, N., Bowlby, M., & Howell, P. (1986). Segmentation and speech perception in relation to reading skill: A developmental analysis. *Journal of Experimental Child Psychology, 41*, 489–507.

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

Steffens, M. L., Eilers, R. E., Gross-Glen, K., & Jallad, B. (1992). Speech perception in adult subjects with familial dyslexia. *Journal of Speech and Hearing Research, 35*, 192–200.

Stothard, S. E., Snowling, M. J., Bishop, D. V. M., Chipchase, B. B., & Kaplan, C. A. (1998). Language impaired preschoolers: a follow-up into adolescence. *Journal of Speech, Language, and Hearing Research, 41*, 407–418.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9*, 182–198.

Tallal, P., & Percy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 12*, 83–94.

Treiman, R., & Breaux, A. (1982). Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults. *Journal of Psycholinguistic Research, 11*, 581–610.

Tunick, R. A., & Pennington, B. F. (2002). The etiological relation between reading disability and phonological disorder. *Annals of Dyslexia: An Interdisciplinary Journal of the International Dyslexia Association, 52*, 75–97.

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.

Walley, A. C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286–350.

Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology, 41*, 48–61.

Wolf, M., & Obregon, M. (1992). Early naming deficits, developmental dyslexia, and a specific deficit hypothesis. *Brain and Language, 42*, 219–247.