What Influences Literacy Outcome in Children With Speech Sound Disorder?

Robin L. Peterson
Bruce F. Pennington
Lawrence D. Shriberg
University of Denver, Denver, CO

Richard Boada
University of Colorado Denver Health Sciences Center

Purpose: In this study, the authors evaluated literacy outcome in children with histories of speech sound disorder (SSD) who were characterized along 2 dimensions: broader language function and persistence of SSD. In previous studies, authors have demonstrated that each dimension relates to literacy but have not disentangled their effects.
Methods: Two groups of children (86 SSD and 37 controls) were recruited at ages 5–6 and were followed longitudinally. The authors report the literacy of children with SSD at ages 7–9, compared with controls and national norms, and relative to language skill and SSD persistence (both measured at age 5–6).
Results: The SSD group demonstrated elevated rates of reading disability. Language skill but not SSD persistence predicted later literacy. However, SSD persistence was associated with phonological awareness impairments. Phonological awareness alone predicted literacy outcome less well than a model that also included syntax and nonverbal IQ.
Conclusions: Results support previous literature findings that SSD history predicts literacy difficulties and that the association is strongest for SSD + language impairment (LI). Magnitude of phonological impairment alone did not determine literacy outcome, as predicted by the core phonological deficit hypothesis. Instead, consistent with a multiple deficit approach, phonological deficits appeared to interact with other cognitive factors in literacy development.

KEY WORDS: speech sound disorders, reading disability, language impairment, literacy, longitudinal

In the present study, we used a longitudinal design to investigate the relationship between childhood speech sound disorder (SSD) and literacy development. It is well documented that as a group, preschoolers with SSD are at increased risk for later reading disability (RD), or developmental dyslexia (Bird, Bishop, & Freeman, 1995; Larrivee & Catts, 1999; Lewis & Freebairn, 1992). Similarly, children with RD have elevated rates of previous SSD (Pennington & Lefly, 2001; Scarborough, 1990). The two disorders are comorbid at a rate of approximately 25%–30% (Gallagher, Frith, & Snowling, 2000; Lewis, 1996; Pennington & Lefly, 2001). The overlap of RD and SSD is not surprising because research during the past several decades has demonstrated that RD is associated with impairments in phonological processing (Vellutino, Fletcher, Snowling, & Scanlon, 2004), whereas the defining symptom in SSD is difficulty producing developmentally appropriate speech. Thus, a common assumption has been that a phonological processing deficit confers risk for both RD and SSD.

Research findings, in fact, indicate more complex associations between SSD and RD. One consistent finding has been that children whose

early language difficulties are limited to expressive phonology have fairly good literacy outcome, whereas children with broader (i.e., syntactic–semantic) language difficulties are much more likely to develop RD (Bishop & Adams, 1990; Catts, 1993; Levi, Capozzi, Fabrizi, & Sechi, 1982; Shriberg & Kwiatkowski, 1988). The overlap of SSD and RD, therefore, appears to be due at least partly to the fact that each is also comorbid with language impairment (LI).

Literature Review

Children with RD have difficulties with accurate and/or fluent word recognition and spelling as well as secondary difficulties in reading comprehension (Lyon, Shaywitz, & Shaywitz, 2003). The most influential explanation of RD in recent decades has been the strong *core phonological deficit hypothesis* (e.g., Stanovich, 1988). According to this hypothesis, a deficit in phonological processing is necessary and sufficient to cause RD, and the severity of a child's reading and spelling problems relates directly to the severity of the underlying phonological deficit. In contrast to this single deficit explanation for RD, a number of investigators have recently proposed a *multiple deficit account* (Pennington, 2006; Snowling, Gallagher, & Frith, 2003). The multiple deficit view holds that a phonological processing deficit puts a child at risk for RD but that this deficit interacts with other risk and protective factors to determine eventual reading and spelling skill. The primary theoretical contribution of the present study is to evaluate whether the comorbidity between SSD and RD is better accounted for by the core phonological deficit account or the multiple deficit account.

A number of early studies investigated academic outcomes in heterogeneous groups of school-aged children or adults with previous histories of various speech-language difficulties. As a group, individuals with early speech-language problems lagged behind their peers in reading, writing, and other academic subjects (Aram & Nation, 1980; Hall & Tomblin, 1978; King, Jones, & Lasky, 1982). A prediction that follows from the core phonological deficit account is that early isolated speech production problems specifically should place a child at risk for RD, whereas broader language impairment might lead to wider academic difficulties. As indicated, however, the research evidence has generally not supported this hypothesis. For example, Bishop and Adams (1990) followed a group of children who had originally been identified at age 4 as having a speech or language problem (Bishop & Edmundson, 1987). At age 8, children who had had isolated expressive phonological difficulties did not differ from controls on word recognition or other literacy measures. In contrast, children with broader language impairments were impaired at word recognition, spelling, and reading comprehension. When these participants were assessed at age 15, a similar pattern emerged on literacy tests, although those with isolated SSD now underperformed controls on phonological awareness and phonological memory tasks (Snowling, Bishop, & Stothard, 2000). Catts (1993) reported that at the end of second grade, children with a history of SSD + LI but not SSD alone performed significantly worse than controls on reading and spelling tasks. Comparison of children with SSD + LI to children with isolated SSD has consistently found much poorer literacy outcomes for those with comorbid LI, even on word-level reading and spelling tasks (Larrivee & Catts, 1999; Leitao & Fletcher, 2004; Leitao, Hogben, & Fletcher, 1997; Lewis, Freebairn, & Taylor, 2000; Nathan, Stackhouse, Goulandris, & Snowling, 2004).

Thus, there is convergent evidence that language status is an important moderator of literacy outcome for children with SSD. Findings have been less consistent on whether isolated SSD confers elevated risk for reading difficulty. Some studies have reported that children with isolated speech problems performed similarly to controls on reading tasks (Bishop & Adams, 1990; Catts, 1993; Nathan et al., 2004; Snowling et al., 2000), whereas others have found evidence of impairment relative to matched controls or national norms (Bird et al., 1995; Leitao & Fletcher, 2004; Lewis & Freebairn, 1992). There are several possible methodological explanations for the inconsistent findings. Some studies had small sample sizes and may have had insufficient power to detect group mean differences. For example, in the Snowling et al. (2000) follow-up of the Bishop and Edmunson (1987) cohort, the subgroup with a history of isolated expressive phonological difficulties included only 10 participants. This group performed numerically worse than controls on a word recognition task, with a nonsignificant effect size (Cohen's d) of 0.44. Participants did perform significantly worse than controls on phonological processing tasks that correlate with reading skill. However, it is unlikely that reduced power accounts for all the negative findings. In the Catts (1993) study, for example, word recognition scores for SSD only and control groups were similar (Cohen's $d = 0.14$).

Other explanations for the mixed findings concern heterogeneity within SSD. Bishop and Adams (1990) proposed the *critical age hypothesis,* which posits that children with isolated SSD will have poor literacy outcome only if their speech difficulties persist to the age when they begin learning to read. Because 75% of children with SSD have normalized speech by age 6 (Shriberg, 1994), this hypothesis helps account for why many children with early phonological impairments do not later meet criteria for RD. Nathan et al. (2004) found some

support for the critical age hypothesis: Children with persisting speech problems at age 6;9 (years;months) performed worse than controls on reading and spelling measures. Support is diminished, however, by the fact that the group with persistent SSD included individuals both with and without LI.

To summarize, literacy development in children with SSD appears to be influenced by a number of linguistic factors, including broader language development and persistence of the speech problem itself. Although a number of longitudinal studies have carefully examined the influence of either comorbid LI or SSD persistence on literacy, no study to date has utilized a design that has allowed these two factors to be fully disentangled. For example, the long-term follow-ups of the Bishop and Edmundson (1987) cohort (Bishop & Adams, 1990; Snowling et al., 2000; Stothard, Snowling, Bishop, Chipchase, & Kaplan, 1998) have extensively compared literacy outcome in children with early isolated speech difficulties versus broader language difficulties but have not accounted for the effect of the persistence of the speech difficulty among children with isolated SSD. Similarly, as mentioned previously, a recent test of the critical age hypothesis (Nathan et al., 2004) did not account for comorbid LI in evaluating the effects of SSD persistence. Inclusion of both factors is important in an evaluation of whether the core phonological deficit hypothesis or the multiple deficit hypothesis better accounts for the relationship between SSD and RD. Research findings to date are more consistent with the multiple deficit hypothesis because they indicate that children with both a phonological deficit and broader language difficulties have the highest rates of RD. However, an alternate explanation is that LI status may have been confounded with SSD persistence—and, thus, with the severity of the phonological deficit—in some earlier studies. If so, then previous findings may be consistent with the core phonological deficit account.

In addition to the linguistic factors of SSD persistence and LI, evidence is also accruing for the importance of nonverbal IQ (NIQ), a nonlinguistic factor, in literacy development among children with early speech-language impairments. Children with speech-language impairments and low NIQ have poorer literacy outcomes than children with speech-language impairments and normal NIQ (Bird et al., 1995; Bishop & Adams, 1990; Catts, Fey, Tomblin, & Zhang, 2002), even accounting for the severity of speech-language impairment (Catts et al., 2002). This finding is somewhat surprising, given the well-established strong relationship between reading and language disorders (Catts & Kamhi, 2005). The relationship between NIQ and literacy may contribute to the heterogeneity of findings in the SSD literature, with studies differing on whether clinical and control groups were matched on NIQ.

Present Study

The present longitudinal study contributes to the research literature by describing the literacy skills of children with histories of SSD, accounting for both broader language function and persistence of speech disorder. During an earlier phase of the study (Time 1; reported in Raitano, Pennington, Tunick, Boada, & Shriberg, 2004), 5- to 6-year-old children with histories of SSD were characterized along two dichotomous dimensions: presence or absence of LI, and persistent or normalized SSD. This classification procedure resulted in four subgroups: (a) normalized SSD no LI, (b) persistent SSD no LI, (c) normalized SSD LI, and (d) persistent SSD LI. Raitano et al. (2004) investigated the pre-literacy skills of the SSD group as a whole and the four subgroups at Time 1. At ages 5–6, the SSD group as a whole performed worse than controls on three robust predictors of later literacy: phonological awareness (PA), letter knowledge, and rapid serial naming (RSN). Consistent with previous literature, predictive associations were stronger for LI status than for SSD persistence status. The two factors were not confounded. There were additive main effects of LI and persistence on PA, and main effects of LI (but not persistence) on RSN and letter knowledge. For RSN, there was also an LI × Persistence interaction reflecting the counterintuitive finding that the subgroup with LI and with normalized speech performed most poorly. Although this finding was unexpected, it was consistent with the fact that this subgroup had the lowest NIQ of all the SSD subgroups. Finally, even the SSD subgroup without LI and with normalized speech performed more poorly than controls on PA tasks.

In the present study, the authors assessed the literacy outcome of these children at ages 7–9 years (Time 2). Instead of relying on dichotomous classifications, primary analyses in the present study treated broader language function and SSD persistence as continuous variables for several reasons. First, the field is increasingly adopting a dimensional rather than a categorical view of language-based learning disorders (Bishop, 1997; Plomin & Kovas, 2005). Second, attrition over the course of the study produced low sample sizes in some of the subgroups; thus, a continuous approach was more methodologically sound. Secondary analyses used dichotomous predictors in order to parallel the published Time 1 results. Because the present study was designed to investigate longitudinal prediction of literacy outcome, all analyses utilized Time 1 language and speech predictor variables. Thus, the term *SSD persistence* refers to the degree to which participants' speech difficulties persisted at age 5–6.

We made the following predictions for literacy outcome in this sample:

Hypothesis 1. As a group, children with histories of SSD will have poorer literacy outcome and higher rates of RD compared with controls.

Hypothesis 2. Among children with SSD, broader language function will predict literacy outcome more powerfully than will SSD persistence.

Hypothesis 3. Prediction of Time 2 literacy skills will be significantly stronger when obtained from a model that includes multiple Time 1 cognitive variables (e.g., PA, broad language, NIQ) than from a model that includes only Time 1 PA.

Hypotheses 1 and 2 follow from previous research. If a substantial portion of individuals with SSD develop RD, then Hypothesis 2 bears on core phonological deficit versus multiple deficit accounts of RD. A finding that children with both a phonological impairment and broader language difficulties have the highest rates of RD would provide strong support for the multiple deficit view, particularly because LI status was not confounded with SSD persistence in this sample at Time 1. Conversely, a finding that SSD persistence is the primary predictor of RD status would provide support for the core phonological deficit view. Hypothesis 3 directly tests whether PA alone can predict—as strongly as multiple cognitive variables—the literacy outcomes of participants in the present longitudinal study.

Method

Participants

One hundred twenty-three 7- to 9-year-old children participated in this study. These children were part of a larger longitudinal and genetic linkage study conducted at the University of Denver examining the relation between SSD and RD. Two groups were recruited for this study: children with a history of childhood SSD (n = 86) and children with no history of speech or language disorder (controls; n = 37). At original assessment (Time 1), children were 5–6 years of age (M = 5;8 years). All children with SSD had a history of speech difficulties by parent report. In addition, SSD participants were required to have a history of speech therapy for speech sound problems or a score at study entry at or below the 30th percentile on a standardized articulation test. Control children were required to have (a) no history of a speech-language disorder by parent report, (b) no history of speech-language therapy, and (c) a score at study entry above the 30th percentile on a standardized articulation test. The average articulation score for the control group at study entry was at the 68th percentile (range: 31st–99th percentile), indicating that all controls earned articulation scores in the average or above-average range. SSD and control participants were recruited through mass mailings to public and private schools in the Denver area as well as through radio and newspaper advertisements. A small number of control participants were recruited through the University of Denver Department of Psychology developmental subject pool. Overall, the goal was to recruit SSD and control participants of similar demographic backgrounds who lived in the same neighborhoods and attended the same schools. Exclusionary criteria for both the SSD and control groups included a known genetic disorder, intellectual disability, a pervasive developmental disorder, significant birth complications, an acquired brain injury, hearing loss, deficits in the peripheral speech articulators (e.g., cleft palate), or a language other than English spoken in the home.

Three participants were excluded from the current Time 2 analyses because they were found not to have met original inclusionary or exclusionary criteria. Of these three participants, 1 control participant had a positive early history of articulation problems, 1 SSD participant was diagnosed with a genetic disorder, and another SSD participant had suspected intellectual disability. In addition, 16 participants (13 SSD and 3 control), or 11.5% of the sample, did not complete the Time 2 visit because their families could not be contacted or declined to participate.

We compared the 13 SSD participants who failed to complete Time 2 with the 86 SSD participants who did complete Time 2 (see Table 1). The two groups did not differ significantly at Time 1 on age, gender, NIQ, or ethnicity. The two groups also did not differ significantly at Time 1 on any speech or language symptom measure nor on any pre-literacy variable. However, the groups differed on the Hollingshead Four Factor Index (Hollingshead, 1975), a measure of socioeconomic status (SES), with participants who completed Time 2 coming from families of higher SES than participants who did not complete Time 2, $t(90) = 2.25$, $p < .05$. These analyses indicate that, with the exception of SES status, the group that completed Time 2 measures is generally representative of the SSD sample that began the study. The number of control participants who did not participate at Time 2 was too small to allow meaningful statistical analyses.

Table 1. Means (and standard deviations) on demographic variables for control and speech sound disorder (SSD) groups.

Variable	Control *M (SD)*	SSD *M (SD)*
N	37	86
Age in months	99.11 (5.92)	98.97 (6.92)
Male (%)	59.3	70.3
Caucasian (%)	72.1	78.4
NIQ	112.00 (7.83)***	104.76 (11.40)
SES	57.37 (7.71)*	54.78 (9.27)

Note. NIQ = Differential Ability Scales Nonverbal IQ Composite; SES = socioeconomic status, as measured by the Hollingshead score.

Control versus SSD group differences: *$p < .05$. ***$p < .001$.

As shown in Table 1, the final control and SSD groups were similar in age, gender, and ethnicity, but there were significant between-group differences on NIQ, $t(97.5) = 4.05$, $p < .001$, and SES (as measured by the Hollingshead Index), $t(117) = 1.99$, $p < .05$, with the control group scoring higher than the SSD participants in both domains. Previous studies have also reported lower NIQ in SSD and LI samples compared with controls (e.g., Johnston, 1994). The difference in the Hollingshead Index was evident despite attempts to match control and SSD participants on SES. A possible explanation for this finding is that because both SSD and LI are familial (Beitchman, Hood, & Inglis, 1992; Lewis, Ekelman, & Aram, 1989; Lewis et al., 2006; Neils & Aram, 1986), parents of children with SSD are at higher risk for language and reading difficulties and, thus, have lower levels of educational and occupational achievement. Because both NIQ and SES are closely related to the SSD phenotype, we followed the same strategy as that used in the Time 1 analyses of this sample. In the primary analyses, we did not covary either variable, but we performed follow-up analyses in which they were covaried to examine their potential association with literacy outcome. In general, assumptions for analysis of covariance (ANCOVA) were met, except that, as just described, there was a significant relationship between our independent variable (SSD vs. control) and the covariates (NIQ, SES). Although use of ANCOVA under such circumstances is controversial, it is widespread in our field, and given the impossibility of random assignment to SSD and control groups, this procedure offered one means of exploring potentially important third-variable issues.

Procedure and Measures

At both Time 1 and Time 2, participants were tested individually at the University of Denver over three sessions lasting 2 hr each. Informed consent was obtained from participants' parents, and assent was obtained from participants over age 7. Tasks were administered in a standardized order across participants. Based on pilot data, a test order was designed to maintain children's interest and engagement throughout a long test battery. In addition, because of concerns about attrition, we made efforts to include the most critical measures (i.e., speech data) on earlier days. Specific test order was somewhat different across the two time points and yet the pattern of results was similar, suggesting that order effects are unlikely to account for the important findings. Table 2 summarizes the constructs assessed at each time point.

Speech production. At both time points, two speech tasks were administered to assess participants' speech production skill: a conversational speech sample used as input to the Speech Disorders Classification System (SDCS; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997a) and the Goldman–Fristoe Test of Articulation (Goldman & Fristoe, 1986). The SDCS is a validated classification instrument used to categorize children (and adults) with speech sound disorder based on a large corpus of lifespan data of individuals with typical and disordered speech. Examiners evoked tape-recorded conversational samples from participants in a naturalistic fashion. The samples were narrowly transcribed using a diacritic system and computer formatting procedures developed specifically for research on child speech disorders (Shriberg, Allen, McSweeny, & Wilson, 2001; Shriberg & Kent, 2003). The SDCS software suite provides both a categorical metric (persistent vs. normalized speech) and a continuous metric indexing severity of speech delay known as the Percentage of Consonants Correct–Revised (PCC-R; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997b). This metric also provides a *z* score adjusted for children's age and sex (*z*PCC-R). The Goldman–Fristoe Test of Articulation is a single-word elicitation task that provides both a raw score for total number of articulatory errors and an age- and sex-adjusted percentile score.

Table 2. Summary of constructs measured at each time point.

Construct	Time 1	Time 2
Speech production	•	•
Language	•	
Nonverbal IQ	•	
Phonological awareness	•	•
Rapid serial naming	•	
Literacy		•

Consistent with previous research (Shriberg, 1994), the large majority of SSD participants had their speech difficulties resolve by age 7–9. The number of individuals with persistent SSD at Time 2 ($n = 7$) was too small to allow meaningful statistical analyses. Thus, analyses of the relation between SSD persistence at Time 2 and literacy are not reported.

Language. At Time 1, the Picture Vocabulary, Oral Vocabulary, Grammatic Understanding, Sentence Imitation, and Grammatic Completion subtests from the Test of Language Development–Primary: Third Edition (TOLD-P:3; Newcomer & Hammill, 1997) were administered to measure broader language skills. The TOLD-P:3 subtests were combined in various ways to give four composites: Expressive Language, Receptive Language, Semantics, and Syntax. Participants were considered to have comorbid LI if they obtained a standard score of 81 or lower (10th percentile) on any of the four composites. This approach to assigning LI status at Time 1 was a slight modification of the EpiSLI system used in past epidemiological research (Tomblin, Records, & Zhang, 1996).

Nonverbal IQ. Nonverbal IQ was estimated at Time 1 with the Pattern Construction and Matrices subtests of the Differential Ability Scales (Elliott, 1990).

Cognitive correlates of literacy. At Time 1, PA was assessed with the Elision, Blending Words, and Sound Matching subtests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) as well as the Bird and Bishop (1992) rhyme task. At Time 2, the same tests were administered except that the Lindamood Auditory Conceptualization task (Lindamood & Lindamood, 1979) was substituted for Sound Matching (since norms for Sound Matching only go up to age 6;11). At Time 1, the Rapid Color and Object Naming subtests of the CTOPP were administered to assess RSN.

Literacy. Four measures of literacy attainment were administered at Time 2. Single-word reading accuracy, single-word spelling accuracy, and reading comprehension were assessed with the Basic Reading, Spelling, and Reading Comprehension subtests of the Wechsler Individual Achievement Test (WIAT; Wechsler, 1992). The Gray Oral Reading Test–Third Edition (GORT-III; Wiederholdt & Bryant, 1992) was used to assess reading accuracy and fluency in connected text.

Data Reduction

Age-corrected standardized scores were used in all analyses. Bird and Bishop's (1992) Rhyme task was the only measure that did not provide such a score; in this case, raw scores were residualized for age. All age-corrected variables had acceptable skewness and kurtosis statistics (absolute value < 3). The number of variables to be used in analyses was reduced by creating composites with subtests designed to assess the same underlying constructs. In some cases (e.g., TOLD-P:3 Language composites, CTOPP Rapid Naming composite), standardized composite scores were provided by the tests employed. When this was not the case, the relevant age-corrected scores were standardized and averaged together.

Results

Predictors of literacy outcome were examined in three ways, corresponding to our three hypotheses. First, we compared rates of categorical RD in the SSD group to rates in controls; we also compared continuously measured literacy outcome in the two groups. Next, to evaluate whether broader language status predicted literacy more powerfully than SSD persistence, we used T1 speech and language variables to predict T2 literacy outcome in a regression model. Follow-up analyses compared literacy outcome across categorically defined SSD subgroups. Third, to more directly test the predictions of the multiple deficit and core phonological deficit accounts of RD, we ran a series of regression equations using T1 candidate deficits to predict T2 literacy.

Hypothesis 1: As a Group, Children With Histories of SSD Will Have Poorer Literacy Outcome and Higher Rates of RD Compared With Controls

We first compared continuously measured literacy outcome in the SSD and control groups. To maximize reliability, we created a composite literacy score from the age-corrected scores on the four literacy tests. The SSD group as a whole performed worse than the control group on the composite literacy variable, $t(121) = 5.11, p < .001$. This finding remained robust when NIQ was covaried, $F(1, 119) = 14.01, p < .001$, and the effect of the covariate was significant, $F(1, 119) = 27.82, p < .001$. The group difference also remained significant when both NIQ and SES (Hollingshead Index) were covaried, $F(1, 114) = 10.25, p < .01$, with significant effects of both covariates: NIQ, $F(1, 114) = 27.60, p < .001$; SES, $F(1, 114) = 11.47, p < .01$.

We next compared rates of categorical RD in the SSD and control groups. RD was defined as a standard score on the GORT-III Fluency composite of 80 or lower (i.e., 9th percentile or poorer). Because the GORT-III uses connected text and measures both speed and accuracy, this definition of RD corresponds well to current, research-validated definitions (Shaywitz, 2003). The GORT-III provides standard scores in five-point increments, so this cutoff was chosen to correspond as closely as possible to the cutoff used for LI (standard score of 81, or 10th percentile). Consistent with the continuous analyses, RD was significantly more prevalent in the SSD group than in the control group, $\chi^2(1, N = 123) = 5.09, p < .05$). Of the SSD sample, 22.1% developed RD, compared with 5.4% of controls. These results are summarized in Table 3.

Hypothesis 2: Among Children With SSD, Broader Language Function Will Predict Literacy Outcome More Powerfully Than Will SSD Persistence

To test this hypothesis, we ran a series of regression analyses within the SSD group. In every case, the dependent measure was the Time 2 composite literacy variable. Each model included two independent variables: the Time 1 composite speech variable (indexing persistence of speech difficulty at study entry) and one of the four Time 1 language composite variables provided by the TOLD-P:3.

Results, which are summarized in Table 4, were very clear. Each of the four language composites predicted later literacy more powerfully than speech. In every

Table 3. Means (and standard deviations) on literacy and cognitive variables for control and SSD groups.

Variable	Control *M (SD)*	SSD *M (SD)*
% RD	5.4*	22.1
Literacy composite *z* score	0.64 (1.03)***	–0.27 (0.86)
WIAT Basic Reading SS	114.16 (17.15)***	100.84 (13.48)
WIAT Spelling SS	111.93 (15.77)***	98.94 (11.51)
WIAT Reading Comprehension SS	115.19 (13.29)***	101.33 (14.48)
GORT-III Fluency SS	108.92 (18.71)***	95.64 (15.60)
T2 Phonological Awareness composite	0.54 (.72)***	–0.23 (1.02)

Note. RD = reading disability; WIAT = Wechsler Individual Achievement Test; GORT-III = Gray Oral Reading Test–Third Edition; T2 = Time 2.

Control versus SSD group differences: *$p < .05$. ***$p < .001$.

case, the association between earlier persistence of speech errors and later literacy was nonsignificant when broader language function was accounted for. However, the association between earlier language skill and later literacy was in the moderate to large range, even after accounting for earlier speech, and the β-weight for each language composite was statistically significant (range: .35–.59; all p values < .01).

These analyses treating T1 speech and language status as dimensional constructs were followed up with analyses treating T1 SSD persistence and T1 comorbid LI as categorical diagnoses. SSD persistence was determined with the dichotomous variable produced by the *z*PCCR, and LI status was determined using a modification of the EPI-SLI system, as described in the Method section. Sample sizes in the resulting four subgroups were as follows: normalized SSD no LI (n = 45), persistent SSD no LI (n = 23), normalized SSD LI (n = 11), and persistent SSD LI (n = 7). Within the SSD group, a 2 × 2 analysis of variance (ANOVA) of the composite literacy variable yielded a main effect of T1 LI status, $F(1, 82) = 21.38$, $p < .001$, indicating that children with SSD + LI had poorer literacy scores than children with isolated SSD. Consistent with the regression results mentioned previously, there was no main effect of T1 SSD persistence ($F < 1$). The same pattern of results held when the outcome variable was categorical RD. Within the SSD group, 66.7% of children with comorbid LI developed RD, compared with 10.3% of those without LI, $\chi^2(1, N = 86) = 26.29$, $p < .001$. In contrast, 26.7% of children with persistent SSD developed RD compared with 19.6% of those with normalized SSD, and this difference was not statistically significant ($p > .4$).

The failure to find an association between persistence of earlier speech difficulties and later literacy after accounting for broader language skill was surprising, given that at Time 1, there were additive main effects of LI status and SSD persistence on PA. To further understand the development of phonological difficulties in our sample, we ran a series of analyses examining the contributions of earlier speech and language skill to T2 PA. Overall, results paralleled those from Time 1. The SSD group as a whole performed worse than controls on the T2 PA composite, $t(95.2) = 4.80$, $p < .001$. The unequal variance procedure was used because Levene's test indicated that the variance in the SSD group was significantly greater than the variance in the control group (.61 vs. .30). Although ANCOVA assumes equal variances, because the larger variance was associated with the group that was larger in size, this violation makes ANCOVA a conservative test. Thus, we followed up by covarying first NIQ and then both NIQ and SES, as in previous analyses. The group difference in PA remained significant when NIQ was covaried, $F(1, 119) = 9.38$, $p < .01$, with a significant effect of the covariate, $F(1, 119) = 17.96$, $p < .001$. The group difference also remained significant when both NIQ and Hollingshead Index were covaried, $F(1, 114) = 6.26$, $p < .05$, with significant effects of both covariates: NIQ, $F(1, 114) = 18.10$, $p < .001$; SES, $F(1, 114) = 10.15$, $p < .01$. Group means and *SD*s are reported in Table 3.

We next ran four regression analyses within the SSD group, predicting Time 2 PA in place of Time 2 literacy from the Time 1 composite speech variable and one of the four Time 1 language composite variables provided by the TOLD-P:3. Results, which are summarized in Table 5, suggest a stronger contribution for persistence of speech errors to later PA than to later literacy. Overall, earlier language skill predicted later PA more powerfully than did earlier speech. However, the β-weight for the speech composite variable was statistically significant in one of the four analyses and contributed unique variance at the trend level in an additional two of the four analyses.

We again followed up these dimensional analyses with analyses treating SSD persistence and comorbid LI as categorical diagnoses. Within the SSD group, a 2 × 2 ANOVA of the T2 PA composite yielded main effects of T1 LI status, $F(1, 82) = 15.30$, $p < .001$, and T1 persistence,

Table 4. Standardized β weights and R^2 values for regressions predicting literacy outcome composite from time 1 (T1) speech composite and each TOLD-P:3 language composite within the SSD group.

Independent variables	β (T1 speech)	β (T1 language)	Total R^2
Speech[a], Expressive Language[b]	.04	.54***	.31
Speech, Receptive Language[c]	.15	.35**	.16
Speech, Semantics[d]	.09	.43***	.21
Speech, Syntax[e]	.07	.59***	.37

**$p < .01$. ***$p < .001$.

[a]Time 1 composite of Goldman–Fristoe Test of Articulation and age-corrected percent of consonants correct scores. [b]Time 1 TOLD-P:3 Expressive Language Composite. [c]Time 1 TOLD-P:3 Receptive Language Composite. [d]Time 1 TOLD-P:3 Semantics Composite. [e]Time 1 TOLD-P:3 Syntax Composite.

$F(1, 82) = 4.22$, $p < .05$, on T2 PA, both in the expected direction. Thus, compared with SSD children whose speech delays had normalized by age 5–6, those with persistent SSD had continuing deficits in PA at ages 7–9. SSD persistence was, thus, associated with a continuing phonological impairment but not with later literacy difficulties, a finding that contradicts the predictions of the core phonological deficit hypothesis.

Hypothesis 3: Prediction of Time 2 Literacy Skills Will Be Significantly Stronger When Obtained From a Model That Includes Multiple Time 1 Cognitive Variables Than From a Model That Includes Only Time 1 PA

To more directly test the predictions of the core phonological deficit versus multiple deficit accounts of RD, we next used Time 1 composites, including both symptom measures and cognitive correlates of literacy, to predict Time 2 literacy outcome. Time 1 composites representing candidate deficits included PA, speech, semantics, syntax, RSN, and NIQ. The Expressive and Receptive Language composites from the TOLD-P:3 were not included because the subtests that compose them were already represented in the Semantics and Syntax composites. We did not include T1 letter knowledge as candidate deficit because it represents a measure of written, rather than oral, language. In order to investigate which of these candidate deficits best discriminated between children who did and did not develop RD, we first examined group mean difference effect sizes for the SSD + RD, SSD only, and control groups. Results are shown in Table 6. With the exception of the speech composite, all candidate deficits discriminated the SSD + RD from the SSD only group, with large effect sizes ($d > 1$). The RSN composite appeared closely related to RD status. Although it discriminated the SSD + RD group from the SSD only group and from controls, it was the only composite that did not differ between the two groups that had not developed RD.

Next, we ran a series of exploratory hierarchical regressions within the SSD sample to determine if T2 literacy composite scores could be predicted from all six

Table 5. Standardized β weights and R^2 values for regressions predicting time 2 phonological awareness composite from time 1 speech composite and each TOLD-P:3 language composite within the SSD group.

Independent variables	β (T1 speech)	β (T1 language)	Total R^2
Speech[a], Expressive Language[b]	.13	.48***	.30
Speech, Receptive Language[c]	.24*	.18†	.10
Speech, Semantics[d]	.21†	.26*	.14
Speech, Syntax[e]	.16†	.51***	.32

Note. TOLD-P:3 = Test of Language Development–Primary: Third Edition.

†$p < .1$. *$p < .05$. ***$p < .001$.

[a]Time 1 composite of Goldman–Fristoe Test of Articulation and age-corrected percent of consonants correct scores. [b]Time 1 TOLD-P:3 Expressive Language Composite. [c]Time 1 TOLD-P:3 Receptive Language Composite. [d]Time 1 TOLD-P:3 Semantics Composite. [e]Time 1 TOLD-P:3 Syntax Composite.

Table 6. Group means and group difference effect sizes for SSD + RD, SSD only, and control groups for time 1 candidate deficits.

Candidate deficit	SSD + RD	SSD only	Control	d_1	d_2	d_3
T1 NIQ	94.83 (9.37)[a]	107.63 (10.32)[b]	112.00 (7.83)[c]	1.27	2.06	0.46
T1 PA	–0.83 (1.39)[a]	0.01 (0.63)[b]	0.57 (0.67)[c]	1.39	2.26	0.87
T1 RSN	85.69 (14.60)[a]	100.95 (15.10)[b]	97.39 (12.86)[b]	1.02	0.87	–0.25
T1 Syntax	82.11 (11.95)[a]	99.00 (13.84)[b]	113.49 (9.59)[c]	1.26	3.02	1.18
T1 Semantics	90.53 (9.07)[a]	105.10 (12.31)[b]	116.95 (9.80)[c]	1.26	2.77	1.04
T1 Speech	–0.58 (0.77)[a]	–0.32 (0.75)[a]	0.88 (0.42)[b]	0.34	2.71	1.90

Note. Cells with the same superscript within the same row do not differ at the $p < .05$ level. RSN = rapid serial naming; d_1 = SSD + RD vs. SSD only; d_2 = SSD + RD versus control; d_3 = SSD only versus control.

[a]Time 1 composite of Goldman–Fristoe Test of Articulation and age-corrected percent of consonants correct scores. [b]Time 1 TOLD-P:3 Expressive Language Composite. [c]Time 1 TOLD-P:3 Receptive Language Composite. [d]Time 1 TOLD-P:3 Semantics Composite. [e]Time 1 TOLD-P:3 Syntax Composite.

candidate T1 deficits. Together, the six variables explained 61% of the variance in literacy outcome. Each predictor was entered both first and last, and we compared the relative magnitudes of ΔR^2 statistics. Of all the composites, PA and Syntax were the most powerful predictors of literacy outcome when entered first, with ΔR^2 change values as follows: PA = .36, $p < .001$; Syntax = .36, $p < .001$; NIQ = .25, $p < .001$; RSN = .09, $p < .05$; Semantics = .21, $p < .001$; and Speech = .04, $p < .1$. When entered last, only PA, Syntax, and NIQ explained significant unique variance in the literacy composite, with ΔR^2 change values as follows: PA = .08, $p < .01$; Syntax = .05, $p < .01$; and NIQ = .03, $p < .05$.

On the basis of these results, we tested a model in which T1 PA, Syntax, and NIQ predicted T2 literacy outcome within the SSD group. β values for all three predictors were significant and comparable (PA = .36, $p < .001$; Syntax = .29, $p < .01$; NIQ = .26, $p < .01$). R^2 was .52, indicating that this model explained 9% less variance in literacy outcome than the model with all six predictors. R^2 did not improve with the inclusion of any of the two- or three-way interaction terms. Based on the effect size results, we then added T1 RSN as a fourth predictor. Although R^2 improved to .61, the β-weight for RSN was nonsignificant ($p > .9$). We believe the model that best balances parsimony with explanatory power includes the three Time 1 predictors of PA, Syntax, and NIQ to predict Time 2 literacy. This model explained 19% more variance in literacy outcome than did PA alone. Further, the β-weights for Syntax and NIQ were similar in magnitude to that for PA. These results are, thus, more consistent with a multiple deficit than core phonological deficit account of RD. Table 7 summarizes total R^2 and standardized β-weights for the model with all six predictors, the core phonological deficit model (PA only), and the final multiple deficit model. A similar set of analyses was carried out in the full sample, including controls. Results were similar except that explanatory power was slightly higher for a model that included RSN in place of NIQ.

Discussion

The purpose of the present study was to examine literacy skills in a sample of 7- to 9-year-old children with

Table 7. Summary of regression models predicting literacy outcome from time 1 variables in SSD sample.

β weight	Model 1 (Exploratory; all predictors)	Model 2 (Core phonological deficit model)	Model 3 (Final multiple deficit model)
(T1 NIQ)	.22*	—	.26**
(T1 PA)	.40**	.60***	.36***
(T1 RSN)	.01	—	—
(T1 Syntax)	.35**	—	.29**
(T1 Semantics)	–.05	—	—
(T1 Speech)	.02	—	—
Total R^2	.61	.36	.52

*$p < .05$. **$p < .01$. ***$p < .001$.

preschool histories of SSD who had been characterized along two dimensions of heterogeneity in SSD at age 5–6: broader language function and persistence of speech delay. Although each of these dimensions had been examined in prior studies, no published studies have examined both dimensions simultaneously to estimate individual and interactive associations with reading disability. The major theoretical question addressed by the present study was whether the comorbidity between SSD and RD is better accounted for by a core phonological deficit or a multiple deficit account of RD. According to the core phonological deficit view, children with a history of SSD are more likely to develop RD because the two disorders share a deficit in phonological processing, and variance in literacy outcome among children with SSD relates primarily to the underlying phonological impairment. According to the multiple deficit view, preschoolers with SSD have a phonological deficit that puts them at risk for later RD, but this deficit interacts with other risk and protective factors to determine literacy outcome. We proposed three hypotheses regarding literacy outcome in SSD, based on previous literature as well as the pre-literacy skills of the current sample, measured at ages 5–6 years. In the paragraphs that follow, we discuss the results in terms of these hypotheses and explain how the results bear on the core phonological deficit versus multiple deficit accounts of RD.

Hypothesis 1: Literacy Outcome in Whole SSD Sample

We hypothesized that as a group, children with SSD would have poorer literacy outcome than controls and would have higher rates of categorical RD. This hypothesis was confirmed. The continuous literacy scores of the SSD group were significantly below those of the control group, even when NIQ and SES were covaried. Further, using a 9th percentile cutoff for the general population, nearly one quarter of the SSD sample (22.1%) met criteria for a reading disability, compared with 5.4% of controls. Thus, the relative risk for RD among children with a preschool history of SSD was 2.5 compared with the general population and 4.1 compared with children with no history of speech-language difficulties. However, as in previous studies, literacy problems were by no means universal in children with a preschool history of a speech sound disorder, and many children in the SSD group achieved literacy scores in the average range (Leitao et al., 1997; Nathan et al., 2004; Snowling et al., 2000). In fact, the literacy test means for the SSD group were close to national norms but still below those of a control group that was somewhat above national norms on both literacy tests and nonverbal IQ. The finding of above-average achievement for a volunteer control sample in our geographic region is not unexpected (e.g., DeFries et al., 1997). Because we recruited our control sample to be similar to the SSD group on demographic variables (including school placement), the children with SSD may well be impaired relative to their peers in the classroom setting. Despite efforts to match control and SSD samples on SES, there were group differences on the Hollingshead Index. However, inclusion of the Hollingshead Index as a covariate did not change the result.

Hypothesis 2: Relative Predictive Value of Earlier Language Versus Earlier Persistence of Speech Errors for Later Literacy

We hypothesized that broad language function at age 5–6 would predict literacy outcome at age 7–9 more strongly than would persistence of speech errors at age 5–6. This hypothesis was confirmed. Regression analyses within the SSD group indicated that the association between early broad language and later literacy was in the moderate to large range after accounting for early speech production. In contrast, the association between early speech production and later literacy was nonsignificant once early broad language was accounted for. The same pattern of results held when comorbid LI and SSD persistence were treated as categorical diagnoses rather than dimensional constructs; there was a main effect of LI but not persistence on literacy. Further, two-thirds of children with SSD + LI (66.7%) met criteria for RD, compared with 10.3% of those with isolated SSD. However, SSD persistence did not discriminate children who developed RD from those who did not. The relative risk for RD among children with SSD + LI was 7.4 compared with that for the general population, and 12.4 compared with that for controls. Thus, the present results agree with a growing body of research indicating that a substantial portion of the overlap between SSD and RD is due to the third variable of LI (Bishop & Adams, 1990; Catts, 1993; Larrivee & Catts, 1999; Leitao et al., 1997; Lewis et al., 2000; Nathan et al., 2004; Snowling et al., 2000). This finding is more consistent with a multiple deficit account than a core phonological deficit of RD and confirms that broader language status interacts with a phonological deficit in determining literacy outcome. Regression analyses indicated that not all aspects of broader language are equally important to literacy outcome, with a syntax composite acting as a more powerful predictor than a semantics composite. Possible explanations for this result, which was not predicted, are discussed next.

Time 1 results in this sample indicated additive main effects of LI and SSD persistence on PA. Although we expected the effect of language on later literacy to be stronger than the effect of speech production, we did not expect to find a null association between SSD persistence and literacy. At Time 2, there was evidence for

a continuing effect of persistence on PA. Thus, the PA tasks appear to index ongoing phonological impairments in 7- to 9-year-old children with SSD. This finding is consistent with previous studies reporting a variety of phonological impairments in individuals with a history of isolated SSD (Bird & Bishop, 1992; Snowling, 2000; Snowling et al., 2000; Webster, Plante, & Couvillion, 1997). Given the robust association between PA and literacy skill (Vellutino et al., 2004), it is quite surprising that there was a main effect of persistence on PA but not on literacy scores.

The present finding poses a serious challenge for a strong core phonological deficit hypothesis of RD and is much more consistent with the multiple deficit account (discussed further in the paragraphs that follow). Present results also do not support the critical age hypothesis. Overall, children whose speech difficulties persisted to the initiation of formal literacy instruction (i.e., Time 1) did not have greater literacy problems than children whose speech difficulties had normalized by that point. Nathan et al. (2004) proposed a modification of the critical age hypothesis—that persistent SSD was associated with RD only when accompanied by a deficit in PA. Our results suggest that literacy outcome for children with both persistent SSD and PA deficits should be relatively good when broader language skill is intact.

The present study treated SSD persistence as a unitary construct. However, it is possible that different underlying cognitive liabilities contribute to a surface-level speech difficulty. For example, some researchers have attempted to discriminate children whose SSD relates primarily to an underlying phonological deficit from children whose SSD relates primarily to peripheral articulatory difficulties (e.g., Dodd & McIntosh, 2008). Future research could investigate whether the relative contributions of SSD persistence and language skill to later literacy are different for these subgroups. We predict that the pattern of results would be similar to the present study, considering that few cases of SSD appear to arise entirely from peripheral articulatory weaknesses (Dodd & McIntosh, 2008; see also Goffman, 2005).

Hypothesis 3: Prediction of T2 Literacy From T1 Cognitive Variables

Finally, in the most direct test of the core phonological deficit versus multiple deficit accounts of RD, we used a wider set of T1 cognitive variables, including both symptom measures and cognitive correlates of literacy, to predict literacy outcome. Results were inconsistent with the core phonological deficit account and more consistent with the multiple deficit account. Compared to a model with PA as the sole predictor of literacy outcome in children with SSD, the model including PA, syntax, and NIQ explained substantially more variance. Effect sizes for the three predictors were similar, clustering in the moderate range.

Although we had expected broader language to relate to literacy outcome, we did not expect a stronger relationship for syntax than semantics. Although we have followed the terminology of the TOLD-P:3 in labeling the composites, it is important to note that the Syntax composite includes three measures, each of which taps slightly different skills. The Grammatic Understanding subtest primarily assesses comprehension of sentence-level syntax; Grammatic Completion primarily assesses productive morphology; and performance on Sentence Imitation is likely constrained by multiple factors, including expressive morphology, expressive syntax, and phonological memory. Thus, although a composite of these tests primarily captures morphosyntactic ability, it may also reflect other factors that have an important relation to literacy, such as phonological memory. Of note, each of the individual subtests in the syntax composite correlated highly with the literacy composite ($r \geq .5$).

The declarative/procedural model of language advanced by Ullman and colleagues (Ullman, 2001, 2004) offers one potential theoretical explanation for the strong relationship between the syntax composite and literacy. According to the model, language is subserved by two separate neural systems: the declarative memory system, which supports learning of specific word meanings, and the procedural memory system, which supports learning of grammatical (and other linguistic) rules. These authors have argued that LI results from a deficit specific to the procedural system (Ullman & Pierpont, 2005). Perhaps the process of learning to read and spell relies more heavily on the procedural memory system because of the many rules governing the mappings between print and sound that must be mastered. Because the present study was not designed to address the differential relation of language component processes to literacy, it will be important for future research to replicate our findings and more carefully investigate alternative explanations.

The present results agree with previous findings demonstrating that holding language skill constant, low nonverbal IQ predicts poorer literacy outcome (Catts et al., 2002). This result is somewhat surprising, given that current theories of reading development emphasize linguistic factors. Horn's *Gf-Gc theory of intelligence* offers one framework for interpreting this relationship (Horn & Noll, 1997). According to this theory, fluid intelligence (Gf) supports the ability to acquire new information, the store of which determines crystallized intelligence (Gc). NIQ measures Gf, whereas literacy skill depends on accumulated knowledge about the pronunciations and meanings of printed words and, thus, reflects Gc. So, Gf-Gc theory holds that NIQ should predict the rate of

acquiring skill in reading and spelling and, thus, the level of literacy skill at a particular point in development. A second (not mutually exclusive) explanation for the NIQ–literacy relationship is that nonverbal skills covary with higher level language skills that support reading development but are not measured by traditional tests of semantic–syntactic ability (Catts et al., 2002). Clearly, further research on this issue is needed.

Conclusions and Future Directions

Overall, results from the present study are more consistent with the multiple deficit than the core phonological deficit account of the comorbidity between SSD and RD. We found clear evidence that SSD is associated with persistent phonological deficits and that children with preschool histories of SSD have elevated rates of RD compared with both the general population and our control group. However, many children with SSD histories did not develop RD despite a lasting phonological deficit. In addition to being problematic for the core phonological deficit account of RD, these results also challenge the more general assumption of a causal relationship between PA and learning to read across the full range of individual differences. This assumption had been prominent in the field for decades and has recently been more carefully scrutinized, most thoroughly in a review by Castles and Coltheart (2004). These authors point out that because few studies have carefully accounted for preexisting literacy skills, much of the relationship between PA and literacy may run in the opposite direction—because learning to read both improves PA skill and changes the way individuals approach PA tasks. Our results suggest that SSD and RD might be associated with poor PA performance for at least partly different reasons.

The present study has methodological limitations that should be addressed in future research. Several important comparisons were based on a control group that consistently performed somewhat above national norms. Although this control group is likely representative of the classmates of the present clinical group, it would be valuable to know whether the present results generalize to children from a broader range of socioeconomic and demographic backgrounds. Further, our study design did not allow us to draw conclusions about literacy in SSD subgroups beyond the age of 7–9. In the present study, the pattern of results was essentially identical for measures of word reading accuracy, paragraph reading fluency, spelling, and reading comprehension. Research has demonstrated that at this early stage of literacy acquisition, reading comprehension is best predicted by word recognition skill, but as children grow older, broader language skills play an increasingly important role in reading comprehension (Curtis, 1980; Gough & Walsh, 1991). A future study should investigate whether dimensions of heterogenetity within SSD relate to differing trajectories for word recognition and spelling versus reading comprehension over time.

Acknowledgments

This work was supported by Grant HD049027 from the National Institute of Health and Human Development. The authors are grateful to Rachel Tunick and Nancy Raitano-Lee for their work on the initial phase of this longitudinal study and to Christa Hutaff-Lee, Irina Kaminer, Lauren McGrath, Erin Phinney, Jenni Rosenberg, and Jennifer Yerks for their work collecting the data reported here as well as for their comments on an earlier draft of this article.

References

Aram, D. M., & Nation, J. E. (1980). Preschool language disorders and subsequent language and academic difficulties. *Journal of Communication Disorders, 13,* 159–170.

Beitchman, J. H., Hood, J., & Inglis, A. (1992). Familial transmission of speech and language impairment: A preliminary investigation. *Canadian Journal of Psychiatry, 37,* 151–156.

Bird, J., & Bishop, D. (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, Language, and Hearing Research, 38,* 446–462.

Bishop, D. V. M. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*. Hove, England: Psychology Press.

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders, and reading retardation. *Journal of Child Psychology and Psychiatry, 31,* 1027–1050.

Bishop, D. V., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52,* 156.

Castles, A., & Coltheart, M. (2004). Is there a casual link from phonological awareness to success in learning to read? *Cognition, 91,* 77–111.

Catts, H. W. (1993). The relationship between speech-language impairments and reading disabilities. *Journal of Speech and Hearing Research, 36,* 948–958.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research, 45,* 1142–1157.

Catts, H. W., & Kamhi, A. G. (2005). *The connections between language and reading disabilities*. Mahwah, NJ: Erlbaum.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology, 72,* 656–669.

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, 7–8,* 19.

Dodd, B., & McIntosh, B. (2008). The input processing, cognitive linguistic and oro-motor skills of children with speech difficulty. *International Journal of Speech-Language Pathology, 10,* 169–178.

Elliott, C. D. (1990). *The Differential Ability Scales*. San Antonio, TX: The Psychological Corporation.

Gallagher, A., 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.

Goffman, L. (2005). Assessment and classification: An integrative model of language and motor contributions to phonological development. In A. G. Kamhi & K. E. Pollock (Eds.), *Phonological disorders in children: Clinical decision making in assessment and intervention* (pp. 51–64). Baltimore: Brookes.

Goldman, R., & Fristoe, M. (1986). *The Goldman–Fristoe Test of Articulation*. Circle Pines, MN: AGS.

Gough, P. B., & Walsh, M. A. (1991). Chinese, Phoenicians, and the orthographic cipher of English. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 199–209). Hillsdale, NJ: Erlbaum.

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

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53–91). New York: Guilford.

Johnston, J. (1994). Cognitive abilities of children with language impairment. In R. V. Watkins & M. L. Rice (Eds.), *Specific language impairments in children: Communication and language intervention series* (Vol. 4, pp. 107–121). Baltimore: Brookes.

King, R. R., Jones, C., & Lasky, E. (1982). In retrospect: A fifteen-year follow-up report of speech-language-disordered children. *Language, Speech, and Hearing Services in Schools, 13,* 24–32.

Larrivee, L. S., & Catts, H. W. (1999). Early reading achievement in children with expressive phonological disorders. *American Journal of Speech-Language Pathology, 8,* 118–128.

Leitao, S., & Fletcher, J. (2004). Literacy outcomes for students with speech impairment: Long-term follow-up. *International Journal of Language and Communication Disorders, 39,* 245–256.

Leitao, S., Hogben, J., & Fletcher, J. (1997). Phonological processing skills in speech and language impaired children. *European Journal of Disorders of Communication, 32,* 73–93.

Levi, G., Capozzi, F., Fabrizi, A., & Sechi, E. (1982). Language disorders and prognosis for reading disabilities in developmental age. *Perceptual and Motor Skills, 54,* 1119–1122.

Lewis, B. A. (1996, November). *The comorbidity of phonology and reading disorders*. Paper presented at the Annual Convention of the American Speech and Hearing Association, Atlanta, GA.

Lewis, B., Ekelman, B., & Aram, D. (1989). A family study of severe phonological disorders. *Journal of Speech and Hearing Research, 32,* 713–724.

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

Lewis, B. A., Freebairn, L. A., & Taylor, H. G. (2000). Follow-up of children with early expressive phonology disorders. *Journal of Learning Disabilities, 33,* 433–444.

Lewis, B. A., Shriberg, L. D., Freebairn, L. A., Hansen, A. J., Stein, C. M., Taylor, H. G., & Iyengar, S. K. (2006). The genetic bases of speech sound disorders: Evidence from spoken and written language. *Journal of Speech, Language, and Hearing Research, 49,* 1294–1312.

Lindamood, C. H., & Lindamood, P. C. (1979). *Lindamood Auditory Conceptualization Test*. Allen, TX: DLM Teaching Resources.

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53,* 1–14.

McGrath, L., Hutaff-Lee, C., Scott, A., Boada, R., Shriberg, L. D., & Pennington, B. F. (2008). Children with comorbid speech sound disorder and specific language impairment are at increased risk for attention deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology, 36,* 151–163.

Nathan, L., Stackhouse, J., Goulandris, N., & Snowling, M. J. (2004). The development of early literacy skills among children with speech difficulties: A test of the "critical age hypothesis." *Journal of Speech, Language, and Hearing Research, 47,* 377–391.

Neils, J., & Aram, D. M. (1986). Family history of children with developmental language disorders. *Perceptual and Motor Skills, 63,* 655–658.

Newcomer, P. L., & Hammill, D. D. (1997). *Test of Language Development–Primary: Third Edition*. Austin, TX: Pro-Ed.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101,* 385–413.

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

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131,* 592–617.

Raitano, N. A., Pennington, B. F., Tunick, R. A., Boada, R., & Shriberg, L. D. (2004). Pre-literacy skills of subgroups of children with speech sound disorders. *Journal of Child Psychology and Psychiatry, 45,* 821–835.

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

Shaywitz, S. (2003). *Overcoming dyslexia: A new and complete science-based program for reading problems at any level*. New York: Alfred A. Knopf.

Shriberg, L. D. (1994). Five subtypes of developmental phonological disorders. *Clinics in Communication Disorders, 4,* 38–53.

Shriberg, L. D., Allen, C. T., McSweeny, J. L., & Wilson, D. L. (2001). PEPPER: Programs to Examine Phonetic and Phonologic Evaluation Records [Computer software]. Madison, WI: Waisman Center.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997a). The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research, 40,* 723–740.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997b). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40,* 708–722.

Shriberg, L. D., & Kent, R. D. (2003). *Clinical phonetics* (3rd ed.). Boston: Allyn & Bacon.

Shriberg, L. D., & Kwiatkowski, J. (1988). A follow-up study of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 53,* 144–155.

Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford, England: Blackwell.

Snowling, M. J., Bishop, D. V. M., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry, 41,* 587–600.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development, 74,* 358–373.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities, 21,* 590–604.

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*(2), 407–418.

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39,* 1284–1294.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30,* 37–69.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92,* 231–270.

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex, 41,* 399–433.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45,* 2–40.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.

Webster, P. E., Plante, A. S., & Couvillion, L. M. (1997). Phonologic impairment and prereading: Update on a longitudinal study. *Journal of Learning Disabilities, 30,* 365–375.

Wechsler, D. (1992). *Wechsler Individual Achievement Test*. San Antonio, TX: The Psychological Corporation.

Wiederholdt, J., & Bryant, B. (1992). *Gray Oral Reading Test–Third Edition*. Odessa, FL: Psychological Assessment Resources.

Received January 28, 2008

Revision received June 20, 2008

Accepted February 9, 2009

DOI: 10.1044/1092-4388(2009/08-0024)

Contact author: Robin L. Peterson, Department of Psychology, University of Denver, 2155 South Race Street, Denver, CO 80208. E-mail: rpeters6@du.edu.