LSHSS

Clinical Forum

The EpiSLI Database: A Publicly Available Database on Speech and Language

J. Bruce Tomblin
University of Iowa, Iowa City

Epidemiology can be defined as the study of the distributions and determinants of health-related states or events in specified populations (Gordis, 1996). Key to this definition is the notion that epidemiology studies health conditions within populations. Rather than performing experiments in which variables are manipulated, epidemiology uses observational or at the most quasi-experimental designs. As a result, it is a science much like astronomy and geology, which look at the world as it is and then exploit naturally occurring patterns that deviate from random and from this infer or hypothesize underlying mechanisms. In the absence of experimental control, it is necessary to use well-designed population sampling methods to obtain control. Sampling biases can create patterns that are not actually in the population. Most crucially, the researcher must believe that all cases within a target population will have an equal probability of being sampled. The populations can consist of clients who have been seen for clinical services, but these populations are often avoided because there are natural biasing factors such as comorbid conditions, severity, help seeking/avoiding behaviors, and access to service that may influence critical features of the condition. In order to conduct such epidemiologic research, it is usually necessary to study large numbers of individuals, and therefore, this type

ABSTRACT: **Purpose:** This article describes a database that was created in the process of conducting a large-scale epidemiologic study of specific language impairment (SLI). As such, this database will be referred to as the EpiSLI database. Children with SLI have unexpected and unexplained difficulties learning and using spoken language. Although there is no uniform standard for the diagnosis of SLI, the construct encompasses a language deficit occurring in the presence of grossly normal sensory and nonverbal cognitive abilities (H. Tager-Flusberg & J. Cooper, 1999). Although these language difficulties are most apparent during the preschool and early school years, evidence now exists that these problems are usually present well into adulthood and are probably present throughout a person's life (see, for instance, C. J. Johnson et al., 1999; S. E. Stothard, M. J. Snowling, D. V. M. Bishop, B. B. Chipchase, & C. A. Kaplan, 1998; J. B. Tomblin, 2008).

Discussion: Much of what we know of these children has come from research on children who have been clinically identified and served. Certainly, by studying those who are being served, our research base is most likely to be relevant to clinical services. However, there is a danger in this research strategy. It is quite possible that not all children with SLI are clinically identified and served within our service delivery systems. In such circumstances, there is the potential for systematic factors to influence which children do or do not find their way to clinical service.
Clinical Implications: If our research questions are concerned with the characteristics of the actual population of children with SLI that exists in our communities and not just those who are being served, then we need to turn to methods of epidemiology to aid our research.

KEY WORDS: public databases, specific language impairment, speech sound disorders, phonological processing

of research is not often done in our field due to its cost. The objectives of the study for which the data were collected were as follows:

- Establish a definition of specific language impairment (SLI) that is consistent with current research and clinical experience, and, in accordance with this definition, develop an explicit criterion for the diagnosis of SLI.
- Estimate, based on this definition, the preva1ence of SLI in children who are entering school within each of three community environments in the United States—urban, suburban, and rural. Furthermore, estimate this prevalence rate for both males and females within these strata and determine the proportion of children with SLI within these strata who had concomitant phonological disorders.
- Evaluate potential risk factors for SLI to determine if any are associated with elevated rates of SLI.
- Determine the history of clinical intervention for those children within each strata who have been identified as having an SLI.

DATABASE OVERVIEW

Background on the Epidemiology of the SLI Database

Approximately 15 years ago, the National Institutes for Deafness and Other Communication Disorders (NIDCD) contracted with the University of Iowa to conduct a study of the epidemiology of SLI. The aims of the study were those described above. This study has generated a large database that was collected to allow us to address these questions (Hammer, Tomblin, Zhang, & Weiss, 2001; Lubker & Tomblin, 1998; Shriberg, Tomblin, & McSweeny, 1999; Tomblin, Hammer, & Zhang, 1998; Tomblin, Records, et al., 1997; Tomblin, Records, & Zhang, 1996; Tomblin & Zhang, 1999; Zhang & Tomblin, 2000). This database is now available to researchers for secondary data analysis purposes. The methods of data collection and a summary of the contents of this database are described below. Additional details of the methods can be found in the publications cited above.

Participants and sampling. The target for this study was a stratified cluster sample of 6,000 kindergarten children who were monolingual speakers. In fact, we sampled 7,218 children. The sample was stratified by residential setting and was clustered according to school building. Rather than choosing a single urban, suburban, and rural area, the sample was drawn from various regions of the states of Iowa and Illinois. These regions were centered on large metropolitan areas that will hereafter be referred to as “population centers.” Each population center was selected for its ability to contribute an urban sample, with the surrounding areas contributing the suburban and rural samples. Several population centers reduced potential bias in participant characteristics associated with a single geographic area. The three selected population centers were Des Moines, IA; Cedar Rapids/Waterloo/Cedar Falls, IA; and the “Quad Cities” that straddle the Mississippi River: Davenport, IA; Bettendorf, IA; Moline, IL; and Rock Island, IL.

Although Iowa is considered overall to be a rural, farming state, the use of population centers provided the desired urban, suburban, and rural residential strata. The selected population centers were the largest in the state. Des Moines, the capital and largest city, had a 1990 metropolitan area population of 338,000. The second largest city in Iowa is Cedar Rapids, with a 1990 metropolitan area population of 170,000. Waterloo/Cedar Falls together had a 1990 metropolitan area population of 147,000. The Quad Cities had a 1990 metropolitan area population of 384,000.

In summary, the population centers selected provided a suitable sample for the study of SLI in monolingual English-speaking children. The general population in the areas sampled provided the linguistic homogeneity desired to reduce the chances that the identified language deficits would be confused with cultural or regional differences.

Strata general definition. The targeted 6,000 kindergarten children were equally distributed into three residential strata: urban, suburban, and rural settings. This stratified sampling was specified by the NIDCD contract and allowed the sampling of children across a spectrum of living and demographic conditions. To achieve this stratified sampling, the attendance zones of the school buildings from the three population centers were drawn and designated as being predominately urban, suburban, or rural. Subsequent to the study, each individual child was assigned to a stratum according to that child’s home address, thus allowing for a more accurate assignment of residential strata.

The U.S. Census Bureau (1995) specifically defines urban and rural areas; however, suburban areas are defined by default, relative to the definitions of urban and rural areas. Based on the U.S. Census Bureau 1995 definitions, “urban” is defined in terms of territory, population, and housing units, and is considered to be places of 2,500 or more persons living in incorporated or unincorporated areas included in urbanized areas. An urbanized area comprises one or more places (“central places”) and the adjacent densely settled surrounding territory (“urban fringe”) that together have a minimum of 50,000 persons. The urban fringe generally consists of contiguous territory having a density of at least 1,000 persons per square mile. The urban fringe also includes outlying territory of such density, connected to the urban area or fringe, and either within 1.5 road miles of the urban core or within 5 road miles of the core but separated by water or other undevelopable territory. “Rural” is defined by the U.S. Census Bureau as territory, population, and housing units not classified as urban. Rural areas may be divided into “places of less than 2,500” and “not in places,” a category that is comprised of rural areas outside incorporated and census designated places and the rural portions of extended cities.

A general rule to determine strata for this study was developed by the investigators based on the two variables of population density and distance from the urban center. Areas designated as being “urban” were within 2 miles of the center business district. “Urban” also included areas that were between 2–3 miles of the center business district if the population density was 3,000 or more people per square mile. “Suburban” designation was assigned to areas having a population density greater than 2,000 persons per square mile and that did not qualify as being urban. “Rural” was considered to be areas with a population density less than 2,000 persons per square mile.

Because of the influence of the Mississippi River on the geographic layout of Rock Island, IL, the following definitions of residential strata for that population center were based solely on population density: Urban was considered to be greater than 3,000 people per square mile, suburban was between 2,000–3,000 people per square mile, and rural was designated as areas having less than 2,000 people per square mile. At study completion, 7,218 children were actually recruited and screened. Table 1 contains the distribution of these participants across the study sites.

Sampling of elementary schools. The method of sampling was a stratified cluster sample of school buildings that were located in the selected population centers. This sampling was accomplished by first contacting in writing the superintendent of each school district in the selected population centers. Along with a written explanation of the study was an invitation to participate during the course of this 2-year study. Receipt of a district superintendent's consent to participate served as permission to contact all of the principals of the school buildings in that district. A total of 41 districts were contacted; 21 (51.22%) superintendents consented to participate, 15 (36.59%) superintendents refused participation, and no response was elicited from 5 (12.19%) superintendents. It should be noted that only public school districts were sampled; there was no sampling of private schools or children being home schooled.

As described above, each participating school building was assigned a residential stratum (urban, suburban, rural) based on its attendance zones. Once the individual school buildings were sorted by population center and residential strata, buildings within each stratum were assigned a number. Using a random number table, buildings were selected to obtain a minimum total sample of 1,000 students in each of the three strata across all population centers. For example, for the testing conducted during Field Year 1, a minimum of 333 children were selected from each rural, suburban, and urban strata in each of the population centers of Des Moines, Waterloo/Cedar Falls, and the Quad Cities. (Cedar Rapids provided us with primarily an urban sample to supplement the urban sample from the Waterloo/Cedar Falls population center).

This procedure was repeated for Field Year 2. Therefore, because of the random sampling, some school buildings did not participate in the study, some were selected to participate in only 1 year of the study, and some were selected to participate in both years of the study. Because the population of Iowa does not contain a substantial number of African Americans, this sampling strategy was modified to oversample the urban strata because this stratum contained the largest proportion of African Americans.

The sampling method described above clearly does not provide for a nationally representative sample; however, the participants who were sampled have been compared with national census data with regard to distribution across residential strata, gender, and ethnicity and were shown to be similar to the census except for a lower proportion of Hispanics and Asians (Tomblin, Records, et al., 1997).

Table 1. Distribution of study participants by strata and study site.

Center	*Rural*	*Suburban*	*Urban*	*Total*
Des Moines	655	789	754	2,198
Cedar Rapids/Waterloo/Cedar Falls	888	665	957	2,510
Quad Cities	814	695	1,001	2,510

SLI EPIDEMIOLOGIC STUDY: MEASUREMENT INSTRUMENTS

Screening

Participants. All kindergarten children in a sampled school were participants in this phase. Parents were informed that this screening was being performed and were given the opportunity to withdraw their child from participation. Negative consent was received for 161 eligible children (2%). The use of this negative consent procedure for this aspect of the study as well as all other aspects of the study was approved by the University of Iowa's Institutional Review Board.

Screening phase screening instrument. The screening procedure only involved language performance. Children were not screened for hearing, nonverbal intelligence, or pervasive developmental disorder, the exclusionary criteria for SLI.

A language screening test was developed that had a very high predictive relationship with the diagnostic outcome (Tomblin et al., 1997). The screening tool consisted of 40 items from the Test of Language Development—Primary: 2 (TOLD–P:2; Newcomer & Hammill, 1988). These items were selected because they were very predictive of the full test score for this age group. This screening instrument was administered to each child individually and took approximately 10 min to complete. It was found to have a test–retest reliability of $r = .80$.

Diagnostic Phase

Participants. All children who failed the screening test and a random sample of those who passed were invited to participate in the diagnostic phase of the study. The control children were sampled from the same school as the screening failures in order to match for residential backgrounds of the SLI cases and controls. Of the 3,877 children who were selected for the diagnostic phase (1,933 screening failures and 1,944 screening passes), 2,084 were given permission by their parents to participate in the study. Of these children, 75 were reported by their parents to speak a second language, which was an exclusionary condition for this study. The rest of the 2,009 children constituted the final monolingual English speaker sample. Of these 2,009 children, 1,929 children provided a full data set.

Diagnostic battery. The goal of the diagnostic testing phase was to identify those children who would serve as SLI cases or control participants. The diagnostic battery included hearing, language, speech, cognitive, and prereading tasks, as well as gross motor observations. Testing was administered in a standardized manner. All children participated individually, and diagnostic testing took approximately 2 hr to complete. The same individuals who administered the screening test also administered the diagnostic tests; however, they were blind to the results of the screen. The diagnostic battery was completely administered during one testing session. When this was not possible due to scheduling reasons, the testing was completed within a week of its initiation. However, individual tests such as the TOLD–P:2 and the Wechsler Preschool and Primary Scale of Intelligence—Revised (WPPSI–R; Wechsler, 1989) were always administered in their entirety during a single session. The order of administration of the diagnostic tests was held constant when possible. Within individual tests, the same presentation order of subtests was maintained across examiners. To supplement the objective measures, the examiners provided written comments regarding the testing situation and their impressions of the child's performance. Additional information concerning these measures can be found in two articles published by the author (Tomblin, Freese, & Records, 1992; Tomblin, Records, et al., 1997).

Audiometric testing. All children in the diagnostic phase were screened for hearing using both pure-tone audiometric screenings and acoustic admittance/impedance audiometry. The latter measures were used to differentiate between persistent hearing loss and loss that might be caused by an ear infection. Pure-tone screening was conducted for 500 kHz, 1 kHz, 2 kHz, and 4 kHz at 20 dB. If a child failed the pure-tone screening bilaterally, no further procedures were done at that time and the child was retested, usually after a period of 2 weeks. If the child failed the pure-tone screening unilaterally, the language diagnostic testing was continued at that time, and unilateral normal hearing was sufficient for a pass.

Language testing. Within the diagnostic phase, standardized measures of listening and speaking were obtained using the TOLD–P:2 and a narrative story task involving narrative comprehension and narrative production (Culatta, Page, & Ellis, 1983). Six TOLD–P:2 subtests were administered: Picture Vocabulary, Oral Vocabulary, Grammatic Understanding, Sentence Imitation, Grammatic Completion, and Word Articulation. The last subtest (Word Articulation) was not used to make the language diagnosis. A complete description of these measures and their scoring can be found in Tomblin, Records, et al. (1997). An estimate of the internal consistency reliability of the items in this language battery was found to be $\alpha = .95$.

Nonverbal IQ. Nonverbal IQ was estimated using the Block Design and Picture Completion subtests of the WPPSI. The reliability of this abbreviated nonverbal IQ test was estimated to be $r = .73$ by retesting 34 children.

Prereading tasks. Early literacy skills were examined as important ancillary measures. The Letter Identification subtest of the Woodcock Reading Mastery Tests—Revised (Woodcock, 1987) and the Word-Sound Deletion task (Catts, 1993) and Random Animals-Colors task (Catts, 1993) were used to measure prereading skills. The Word-Sound Deletion task, which is an elision task, asked children to report what word would result by deleting a morpheme, syllable, or phoneme from a word (What is stop without the /s/?) and thus is a reflection of phonological awareness. The Random Animals-Colors task, which is a rapid automatized naming (RAN) task, asked children to name 24 pictures of a "horse," "cow," and "pig" depicted in three colors (blue, red, or black) that were arranged on a page in rows. This task provided a single score representing the time taken to name the 24 pictures.

Iowa Severity Rating Scale. The Iowa Severity Rating Scale (ISRS: Jeffery & Freilinger, 1986) was developed for use by speech-language pathologists (SLPs) working in the Iowa Public School System. The purpose of the ISRS in this study was to supplement the standardized measures of speech and language with a clinically significant measure and to obtain an informal measure of voice and fluency.

The ISRS is a 5-point severity rating scale of speech, language, voice, and fluency skills. The ratings are on a continuum, where 0 indicates adequate skills and 4 indicates a disorder. There are specified published criteria to guide the rating made by the SLP, and these guidelines were used during this study. Minimal additional guidelines were established for use during the field study because the examiners' subjective impressions were desired. The Word Articulation (WA) subtest of the TOLD–P:2 was supplemented by the ISRS articulation rating that was made by the examiner. A rating of "1" indicated developmental s, r, l problems. Because certain phonemes (initial k, m, v, n) are not sampled by the WA subtest, there was in some cases a discrepancy between the WA results and the articulation severity rating. The language severity rating was determined by the SLP based on observation of the child in informal interactions as well as during the standardized language testing. In some cases, the TOLD–P:2 raw scores were converted to standard scores by the examiners, and this information was considered when making the severity rating for language, as specified by the ISRS manual. For children with low performance scores on the WPPSI, the language severity rating was made commensurate with their cognitive ability. The only guidelines provided to supplement the manual for scoring voice were that mild hypernasality was rated as a "2," and moderate hypernasality was rated as a "3." Ratings concerning fluency were made in accordance with the ISRS manual for fluency.

Motor skills. Two gross measures of motor skills were obtained: gait and handedness. The examiner observed the child walking to the examining room and noted if there were any obvious gross motor problems when walking. The child was also asked to write his or her name, and the examiner made note of which hand the child used.

Diagnostic Data Entry and Diagnosis

The diagnostic data were coded using a data entry program. To minimize coding errors, the data were entered using a two-pass method. Data were first entered and then verified during a second entry of the same data. Thus, data were entered at two different times and usually by two different people. The data entry program also had range checking for data having a minimum low and maximum high value.

Diagnostic outcomes were determined based on language performance as well as hearing and nonverbal cognitive performance using the EpiSLI diagnostic criterion (Tomblin et al., 1996). Based on the diagnostic results, the children were assigned to one of four diagnostic categories: (a) specific language impaired (SLI; failed the language testing but passed the hearing and nonverbal cognitive testing); (b) control (C; passed all language, hearing, and nonverbal cognitive testing); (c) nonspecific language impaired (NLI; failed the language testing and nonverbal cognitive testing, but passed the hearing testing); and (d) cognitive failure (CF; passed the language and hearing testing, but failed the nonverbal cognitive testing). As reported by Tomblin, Records, et al. (1997), 216 of the 1,929 children were identified as SLI, 1,287 children qualified as controls, 209 were identified as NLI, and 222 were identified as CF.

Risk Factor Survey

A major research question in this study was concerned with identifying risk factors for SLI. In order to address this question, information was obtained from the parents of 177 children with SLI and 984 language normal controls regarding a wide range of exposures to potential risk factors. The candidate risk factors were selected after a literature review of risk factors and language impairment. A pilot survey was developed to obtain information regarding these risk factors and was administered to 6 parents.

Several modifications were made based on the pilot. The resulting questionnaire consisted of 180 items that were coded into more than 700 variables.

The Risk Factor Survey included questions about the parents' demographic information, including parental age, marital status at the time of the study, child's birth, race, education, income, and work history during the 12 months before the child's birth. Another section focused on the health history and pre- and postnatal exposures of each parent. This section included questions concerning the parents' medical history; dates and types of infectious diseases; high blood pressure; conditions of the immune system; work-related chemical and physical exposures; and use of alcohol, illicit drugs, and tobacco. Mothers were also asked about adverse prenatal or reproductive conditions and medications used during pregnancy, labor difficulty, delivery type, and birth weight. Finally, the questionnaire asked about features of the child's rearing environment, including breast feeding and parenting practices during the preschool years as well as key developmental milestones and behaviors during this time.

Risk Factor Survey administration. The names of the parents who were to receive the Risk Factor Survey were sent from the laboratory at the University of Iowa to the statistical laboratory at Iowa State University (ISU), a subcontractor on this contract. The lab at the University of Iowa contacted these parents by letter to notify them that they had been selected to receive the Risk Factor Survey. Trained professional interviewers at the ISU statistical laboratory then called the parents to arrange a mutually convenient time to administer the survey. The interviewer usually administered the survey during one call. The parents almost always received the telephone survey before receiving any information regarding the outcome of their child.

A supplementary information booklet was developed to aid the telephone interviewers and to ensure standardized administration of the survey. This booklet explained terms used in the survey as well as general instructions for the interviewer on a question-by-question basis. All interviewers were blind regarding the diagnostic outcome of the study child.

Prior uses of the data. The data obtained from this study have been used to generate a series of articles by the investigators. The principal results of the prevalence study were reported in three articles (Shriberg et al., 1999; Tomblin et al., 1996; Tomblin, Records, et al., 1997). Results of the risk questionnaire study have been summarized in two articles. One concerned the association of prenatal risk exposures with SLI (Tomblin, Smith, & Zhang, 1997); the other examined information concerning the parents' rearing practices and the association of these factors with SLI (Hammer et al., 2001). Additional articles have also used data to evaluate psychometric properties of the TOLD–P:2 (Hammer, Pennock-Roman, Rzasa, & Tomblin, 2002) and to further explore the association of breastfeeding with SLI (Drane, 2003).

Access to the database. The EpiSLI database is available for distribution on request at no charge as part of the National Institutes of Health's initiative for data sharing.[1] A form is provided on the Web site for making this request. Upon agreement with fair use guidelines, individuals will be sent a compact disc containing the data and documentation of the data.

[1]A Web site (http://www.uiowa.edu/~clrc/epidemiologic/index.html) is available that describes the database and provides the instructions for requesting the data.

EXAMPLE OF DATABASE USE: THE RELATIONSHIP OF PRELITERACY SKILLS WITH SPEECH SOUND AND/OR LANGUAGE PROFICIENCY

One of the most prominent explanations for reading disability, and in particular, dyslexia, can be found in the phonological core deficit hypothesis (see, for instance, Liberman, 1973; Snowling & Hulme, 1994; Stanovich, 1988). This theory claims that successful reading development, particularly with an alphabetic language, is dependent on well-developed phonological representations. Recently, Raitano, Pennington, Tunick, Boada, and Shriberg (2004) argued that the phonological core deficit hypothesis should also predict that speech sound disorders (SSD) would be strongly associated with a reading disorder. There are several studies that have shown that children with SSD are at greater risk for a reading disorder (see, for instance, Bishop & Adams, 1990; Larivee & Catts, 1999; Lewis & Freebairn, 1992; Snowling, Bishop, & Stothard, 2000) than children with typical speech development. There are also other studies that suggest that children with SSD are either at no or very low risk for poor reading outcomes during the school years (Beitchman et al., 1996; Bishop & Adams, 1990; Catts, 1993; Hall & Tomblin, 1978; Levi, Capozzi, Fabrizi, & Sechi, 1982). The research on preliteracy skills involving phonological awareness, letter identification, and rapid serial naming in children with SSD shows rather consistent results, indicating that these children have poorer preliteracy skills than children with typical speech sound development (see, for instance, Carroll & Snowling, 2004; Hesketh, Adams, Nightingale, & Hall, 2000; Nathan, Stackhouse, Goulandris, & Snowling, 2004; Raitano et al., 2004; Rvachew & Grawburg, 2006).

These data collectively provide evidence in support of the hypothesis that SSD are a risk factor for a reading disorder. One conclusion that was reached in several studies was that reading outcomes were much worse if SSD were accompanied by language impairment. In fact, in all of the studies that failed to find an elevated risk for reading disorder associated with SSD, the language status of the children with SSD was controlled. Additionally, all of the studies that reported a relationship between SSD and either concurrent preliteracy abilities or subsequent reading used samples that were clinically identified. In contrast, the one study that used a population sample (Beitchman et al., 1996) did not find an association between SSD and later reading. Clinically identified samples are more likely to contain children with comorbid conditions such as concomitant language impairment. None of the studies examining preliteracy in children with SSD was performed with children who were sampled via a population sampling scheme.

The purpose of this example study was to determine whether the findings of previous studies concerning preliteracy skills in children with SSD could be replicated within a population sample. The children in the EpiSLI study were given the Articulation subtest of the TOLD–P:2 and were also administered the Word-Sound Deletion task and the Random Animals-Colors task. Additionally, the children were given the language measures described earlier from which a diagnosis of language impairment was determined based on the EpiSLI criteria described by Tomblin, Records, et al. (1997). We can examine these data to see with a quite large sample drawn from a population sample, whether speech sound ability is associated with phonological awareness (PA) as represented by elision and rapid serial naming as measured by the Random Animals-Colors task. Furthermore, we can examine whether this relationship changes with SSD determined by different levels of severity.

METHOD

Participants

Data from the 1,929 children who were participants in the diagnostic phase of the EpiSLI study were used for this analysis. The children contributing these data averaged 6;1 (years;months) in age. Boys (54.6%) were slightly more common in this sample than girls (45.4%).

Measures

Spoken language status. All children were given the language battery based on the TOLD–P:2 and the narrative story task. LI was determined according to the EpiSLI diagnostic standard described in Tomblin et al. (1996). In this study, performance IQ was not incorporated into the diagnosis; thus, these children were either SLI or NLI.

Speech sound ability. Speech sound performance was assessed using the Articulation subtest of the TOLD–P:2. SSD was assigned to children with scores below the 3rd percentile (2 *SD*s below the mean).

PA. The Word-Sound Deletion task was administered to each child in the diagnostic phase of the study. This task can be viewed as a measure of PA. This task requires the child to delete the initial phoneme or syllable from a word and repeat only the remaining phoneme or syllable. Three example items, all compound words, were demonstrated using pictures. For example, the directions were "Say 'baseball'" as pictures of a "base" and a "ball" were shown. Then, with the first picture covered, the child was asked "Now say 'baseball' without the 'base.'" If the child did not respond correctly, the correct answer was provided for the demonstration items.

When the child showed an understanding of the task, the testing began. The directions were the same for each of the 21 test items; however, the pictured stimuli were discontinued. The stimuli consisted of compound words, two-syllable words, and monosyllabic words. The sound sequence remaining as the correct response was always a high-frequency word. One repetition of an item was provided if needed, and testing was discontinued when six consecutive items were answered incorrectly. Scores on this task were converted to *z* scores based on norms computed from the participants in this study. Poor PA was determined to be performance at or below –1 *SD*.

RAN. The Random Animals-Colors task measured rapid naming ability, a skill that has been reported to be a measure of phonetic coding ability (Catts, 1993). The Random Animals-Colors task involved showing the child an 11.5" × 17.5" page that contained images of 24 animals. These 24 animals were one of three randomly selected animals (a pig, a horse, and a cow) that were colored in one of three randomly selected colors (blue, red, or black). These colored animals were arranged in random order in four rows of six items each. The child was first given as much practice as needed to identify the animals and colors, and several demonstration items were provided to allow the child to practice responding with "adj+noun" responses. If the child did not know his or her colors or animals, this task was not administered. The examiner instructed the child to "name these as fast as you can" in sequential order. The reported score for this task was the total time required for the child to name all of the colored animals, as measured with a stopwatch. Thus, a lower total time score reflects better performance. A tally of incorrect responses was also kept for this task but was not used for this analysis. The naming time in seconds was converted to *z* score values, and naming times that were ≥1 *SD* above the mean were considered to be poor naming times.

RESULTS

This study asked whether children attending kindergarten who had poor speech sound skills were more likely to have poor preliteracy skills than children with typical speech development. Because SSD and SLI have a moderate comorbidity, we also examined the relationship of SSD in children with and without LI and this relationship in children with LI and no SSD.

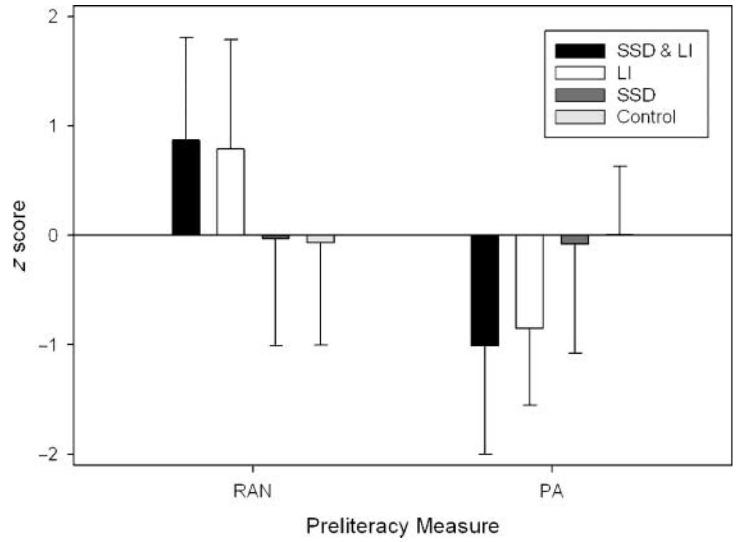
Table 2 contains the rates of poor PA and poor RAN in the children with typical speech sound abilities and those with SSD. These data show a significant association of SSD with PA, $\chi^2(1, N = 1929) = 9.48$, $p < .002$, but not SSD with RAN, $\chi^2(1, N = 1929) = 1.99$, $p = 0.16$. The association found between SSD and PA, however, is not particularly strong. The effect size as represented by a phi coefficient is only .07 and therefore falls in the range of a small effect. It is also important to note that the classification of SSD in this case did not consider the child's language status, which has been shown to be associated with SSD (Shriberg et al., 1999).

The association of SSD with or without LI was examined in order to explore whether the association found above was influenced by the child's language status. Figure 1 provides the means and standard deviations for the PA and RAN tasks for the children with SSD, LI, SSD and LI, as well as the control children. In this figure, the RAN performance represents *z* score values of the log-naming time. Because longer times represent slower performance, higher values on the RAN are reflected in higher positive values in this figure. These data show that performance levels on both preliteracy tasks for the children with LI, regardless of associated SSD, are similar. Likewise, the performance levels of the children with SSD were similar to the control children. Table 3 presents the rates of poor PA and RAN performance in the subgroups of children with both SSD and LI, SSD alone, LI alone, and then the children with typical language and speech status. When the three groups with SSD and/or LI were compared with the typically

Table 2. Rates of poor phonological awareness (PA) and poor rapid automatized naming (RAN) during the Random Animals-Colors task for children with typical speech development and those with speech sound disorders (SSD).

	Poor PA	*Poor RAN*	*Number*
Typical speech	31.38%	17.28%	1,788
SSD (<–2.0 *SD*)	43.97%	21.99%	141

Figure 1. Means (standard deviation) for the Word-Sound Deletion task (phonological awareness; PA) and the Random Animals-Colors task (rapid automatized naming; RAN) for children with speech sound disorders (SSD), language impairment (LI), or both (SSD & LI). The values for the RAN are standardized reaction times; positive values represent slower than average reaction times.



developing children with respect to the rate of poor PA, we found a significant association of PA deficits in the children with SSD plus LI, $\chi^2(1, N = 1466) = 193.20$, $p < .0001$, and the children with LI alone, $\chi^2(1, N = 1466) = 57.59$, $p < .0001$, but no significant association of PA deficits for the children with SSD alone, $\chi^2(1, N = 1509) = 1.58$, $p < 0.21$. The same comparison was performed for the association of RAN with SSD and/or LI. A significant association was found between RAN deficits and SSD plus LI, $\chi^2(1, N = 1466) = 36.90$, $p < 0.001$, as well as between RAN and LI alone, $\chi^2(1, N = 1788) = 155.65$, $p < 0.001$. No significant association was found between SSD and RAN, $\chi^2(1, N = 1509) = 0.01$, $p < 0.91$. In order to aid in the interpretation of these data, it is helpful to present the strength of associations in a way that considers the different sample sizes. In Figure 2, the effect sizes as represented by a phi coefficient are presented. These data show that the strength of association for both PA and RAN is much greater for contrasts between typically developing children and children with LI than contrasts involving children with SSD only. Surprisingly, the largest effect sizes for both RAN and PA were for the contrast of the LI only versus control group. In both cases, the contrast of the children with LI and SSD against the controls yielded smaller effect sizes; however, the difference in rates of poor PA or RAN in the SSD & LI compared with the LI alone were not significant. Thus, SSD does not add significantly to the risk of poor preliteracy skills in children with LI.

Table 3. Rates (percentages) of poor PA and poor RAN during the Random Animals-Colors task as well as the total number of children with SSD and LI, children with SSD only, children with LI only, and children with typical speech and language development (TD).

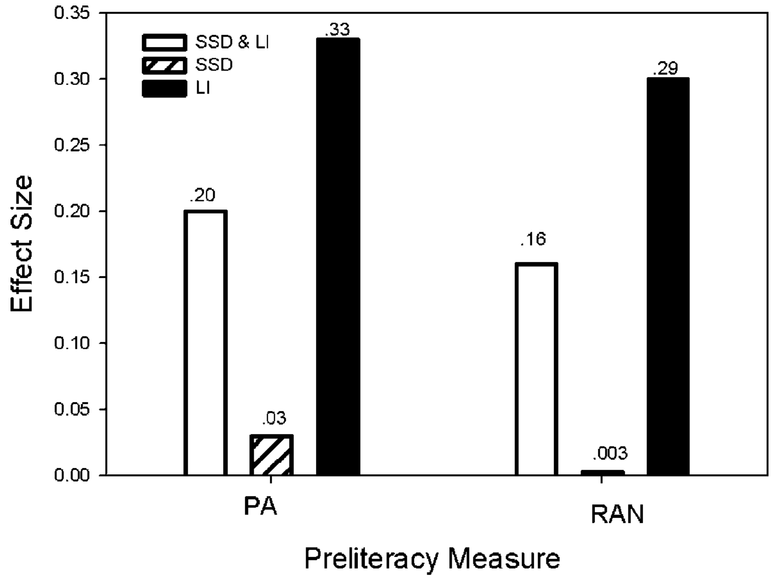
	Poor PA	*Poor RAN*	*Total number*
SSD and LI	71.43%	40.82%	49
SSD only	29.35%	11.96%	92
LI only	61.19%	39.08%	371
TD	23.57%	11.57%	1,417

DISCUSSION

The data from this large sample of children attending kindergarten indicate that there is at best a very weak relationship between SSD and PA as measured by the elision task, and that even this relationship may be the product of the covariance of language with SSD. The PA skills of the children with SSD alone were no different than those of the children with typical speech development. The elision task is similar to many tests of PA, where the least difficult items involve morphological and syllabic structures, and the more difficult items involve phoneme deletion. Many of these kindergarten-age children were not able to progress to the items involving the deletion of phonemes, and thus it remains possible that a special relationship exists between PA and speech sound production that is not found with PA of higher order phonological units.

The data also provided no support for an association between RAN and SSD among children of this age. The fact that no relationship was found might be attributed to the fact that PA and RAN were each measured by just a single measure, and thus these measures were not sufficiently sensitive to reveal a relationship. Although this remains a possibility, these measures were sufficient to reveal a strong relationship with individual differences in language. It was particularly interesting to see that the children with only LI were more likely to have poor preliteracy skills than the children with both SSD and LI. If SSD had its own effect on preliteracy, we would expect the children with both deficits to be

Figure 2. Effect sizes (phi coefficient) for the Word-Sound Deletion task (phonological awareness) and the Random Animals-Colors task (RAN) for children with SSD, LI, or both (SSD & LI).



more likely to have the poorest preliteracy skills. The poor relationship between these early literacy measures and SSD could be due to speech therapy having a positive effect on them and in particular on PA (Hesketh et al., 2000). Many of the activities performed during therapy for an SSD should promote PA and thus would obscure a relationship. The EpiSLI database contains two variables that are concerned with clinical identification and therapy. The parents were asked if their child had been identified as having a speech-language problem; those who said "yes" were asked if the child had received therapy. Of the 1,924 children for whom this information was obtained, 283 (15%) had a history of identification and 204 (11%) had received therapy. Of the 272 children who had SSD alone, LI alone, or SSD and LI, 115 had received therapy according to the parent report. A comparison of PA scores for the treated versus nontreated children with SSD and/or LI showed no difference $t(270) = 0.65, p = 0.51$ in this measure. Therefore, receipt of speech-language services did not seem to influence the findings of this study.

These findings contrast with the existing studies that have examined the prelinguistic skills of children with SSD. One possible reason for this could be that many of these children with SSD had not been identified via community clinical service providers. In the one other study that examined the reading outcomes of children with SSD within a population sample, similar null results were found (Beitchman et al., 1996). This raises the possibility that factors that are associated with clinical identification such as severity or comorbid conditions could account for the heightened risk for poor reading that was found in the clinically served children with SSD.

Conclusions

This study was provided as an example of using the EpiSLI database to largely replicate prior research. In this case, the results did not replicate the earlier findings. Thus, it would appear that further consideration of this issue is warranted. The study just presented exemplifies the type of research that can be conducted with this database. For certain research questions, these data can be useful in allowing an investigator to explore possible relationships among speech and language skills in kindergarten-age children. Clearly, the principal value of this database lies in the number of children used in the study and, as noted earlier, the fact that these children were sampled via population methods. It also has many limitations. In particular, the measures of speech, language, and preliteracy were not extensive: They were selected in order to do a particular job as efficiently as possible within the constraints of measurement reliability and validity. Specifically, the primary measures were concerned with providing an estimate of the prevalence of SLI and the risk factors for SLI; therefore, our inclusion of preliteracy and other associated skills was of secondary importance. In the study just described, it would have been preferable to have had more extensive measures of PA and, for that matter, speech sound production. As seen in the above study, as well as future research needs, these constraints may make it difficult to derive conclusive findings; however, these data may be useful as pilot data or in cases where a replication of findings is needed. As noted in the introduction, many of the primary questions of interest that can be addressed with these data have already been published; however, this does not mean that the data have been exhausted with regard to new insights. I would note that the data from the telephone questionnaire is quite extensive, and although we have examined these data with regard to the primary questions, these data could yield unexpected insights via data mining methods. Along these same lines, the data in this database can be quite useful for teaching purposes, particularly with regard to learning about sampling and multivariate statistics. The data contained in the EpiSLI database were obtained via public funds and represent an investment toward improving our understanding of SLI. The database is being made public in order that we can glean as much out of this investment as possible; toward that end, I hope that future investigators can benefit from it.

ACKNOWLEDGMENTS

The research reported in this article was supported by Contract N01-DC-1-2107, Supplement 3 P50 DC002746-08S1, and Grant P50 DC 2746 (Tomblin, P.I.) from the National Institute on Deafness and Other Communication Disorders, a division of the National Institutes of Health.

REFERENCES

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., Inglis, A., Lancee, W., et al. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 35,* 804–814.

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

Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry, 45,* 631–640.

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

Culatta, B., Page, J., & Ellis, J. (1983). Story retelling as a communicative performance screening tool. *Language, Speech, and Hearing Services in Schools, 14,* 66–74.

Drane, D. L. (2003). *Breast feeding and specific language impairment*. Unpublished doctoral dissertation. Northwestern University, Evanston, IL.

Gordis, L. (1996). *Epidemiology.* Philadelphia: W. B. Saunders.

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.

Hammer, C. S., Pennock-Roman, M., Rzasa, S., & Tomblin, J. B. (2002). An analysis of the Test of Language Development—Primary for item bias. *American Journal of Speech-Language Pathology, 11,* 274–284.

Hammer, C. S., Tomblin, J. B., Zhang, X. Y., & Weiss, A. L. (2001). Relationship between parenting behaviours and specific language impairment in children. *International Journal of Language & Communication Disorders, 36,* 185–205.

Hesketh, A., Adams, C., Nightingale, C., & Hall, R. (2000). Phonological awareness therapy and articulatory training approaches for children with phonological disorders: A comparative outcome study. *International Journal of Language & Communication Disorders, 35,* 337–354.

Jeffery, R., & Freilinger, J. J. (1986). *Iowa's Severity Rating Scales for Speech and Language Impairments*. Danville, IL: Interstate Printers & Publisher.

Johnson, C. J., Beitchman, J., Young, A., Escobar, M. D., Atkinson, L., Wilson, B., et al. (1999). Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research, 42,* 744–760.

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

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

Liberman, I. Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society, 23,* 64–77.

Lubker, B. B., & Tomblin, J. B. (1998). Epidemiology: Informing clinical practice and research on language disorders of children. *Topics in Language Disorders, 19,* 1–26.

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.

Newcomer, P., & Hammill, D. (1988). *Test of Language Development—Primary: 2*. Austin, TX: Pro-Ed.

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.

Rvachew, S., & Grawburg, M. (2006). Correlates of phonological awareness in preschoolers with speech sound disorders. *Journal of Speech, Language, and Hearing Research, 49,* 74–87.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1461–1481.

Snowling, M., 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 and Allied Disciplines, 41,* 587–600.

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London Series B—Biological Sciences, 346,* 21–27.

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,* 407–418.

Tager-Flusberg, H., & Cooper, J. (1999). Present and future possibilities for defining a phenotype for specific language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1275–1278.

Tomblin, J. B. (2008). Adolescent outcomes of SLI. In C. Norbury, J. B. Tomblin, & D. V. M. Bishop (Eds.), *Developmental language impairment* (pp. 93–114). London: Psychology Press.

Tomblin, J. B., Freese, P. R., & Records, N. L. (1992). Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech and Hearing Research, 35,* 832–843.

Tomblin, J. B., Hammer, C. S., & Zhang, X. Y. (1998). The association of parental tobacco use and SLI. *International Journal of Language & Communication Disorders, 33,* 357–368.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40,* 1245–1260.

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.

Tomblin, J. B., Smith, E., & Zhang, X. (1997). Epidemiology of specific language impairment: Prenatal and perinatal risk factors. *Journal of Communication Disorders, 30,* 325–344.

Tomblin, J. B., & Zhang, X. (1999). Are children with SLI a unique group of language learners? In H. Tager-Flusberg (Ed.), *Neurodevelopmental disorders: Contributions to a new framework from the cognitive neurosciences* (pp. 361–382). Cambridge, MA: MIT Press.

U.S. Census Bureau. (1995). *Urban and rural definitions*. Retrieved October 25, 2008, from http://www.census.gov/population/censusdata/urdef.txt.

Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence—Revised*. San Antonio, TX: The Psychological Corporation.

Woodcock, R. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: AGS.

Zhang, X., & Tomblin, J. B. (2000). Factors related to receipt of speech-language therapy. *American Journal of Speech-Language Pathology, 9,* 345–357.

Received May 23, 2008
Accepted February 10, 2009
DOI: 10.1044/0161-1461(2009/08-0057)

Contact author: J. Bruce Tomblin, University of Iowa, Department of Speech Pathology and Audiology, 250 Hawkins Drive, 7 WJSHC, Iowa City, IA 52242. E-mail: j-tomblin@uiowa.edu.