
Prevalence of Speech Delay in 6-Year-Old Children and Comorbidity With Language Impairment

Lawrence D. Shriberg
University of Wisconsin
Madison

J. Bruce Tomblin
University of Iowa
Iowa City

Jane L. McSweeney
University of Wisconsin
Madison

We estimate the prevalence of *speech delay* (L. D. Shriberg, D. Austin, B. A. Lewis, J. L. McSweeney, & D. L. Wilson, 1997b) in the United States on the basis of findings from a demographically representative population subsample of 1,328 monolingual English-speaking 6-year-old children. All children's speech and language had been previously assessed in the "Epidemiology of Specific Language Impairment" project (see J. B. Tomblin et al., 1997), which screened 7,218 children in stratified cluster samples within 3 population centers in the upper Midwest. To assess articulation, the Word Articulation subtest of the Test of Language Development-2: Primary (Newcomer & Hammill, 1988) was administered to each of the 1,328 children, and conversational speech samples were obtained for a subsample of 303 (23%) children. The 6 primary findings are as follows: (a) The prevalence of speech delay in 6-year-old children was 3.8%; (b) speech delay was approximately 1.5 times more prevalent in boys (4.5%) than girls (3.1%); (c) cross-tabulations by sex, residential strata, and racial/cultural backgrounds yielded prevalence rates for speech delay ranging from 0% to approximately 9%; (d) comorbidity of speech delay and language impairment was 1.3%, 0.51% with Specific Language Impairment (SLI); (e) approximately 11–15% of children with persisting speech delay had SLI; and (f) approximately 5–8% of children with persisting SLI had speech delay. Discussion includes implications of findings for speech-language phenotyping in genetics studies.

KEY WORDS: articulation, disorders, epidemiology, genetics, phonology

The prevalence of a public health problem is a key statistic in the allocation of funds for research and treatment. Prevalence estimates also play a central role in behavioral and molecular genetic techniques that require estimates of the risk or liability for the target disorder. Liability estimates are used in analyses that test the fit of findings to possible ways the genotype is transmitted, such as major locus, sex-linked, or multifactorial transmission modes. Moreover, when oligogenic, polygenic, or multifactorial modes of inheritance are suspected, valid prevalence data can be used to estimate the number of genes that may be contributing to the behavioral phenotype.

There currently are no reliable estimates of the prevalence of different forms of child speech disorders for emerging genetic studies and other research and applied needs. Relevant methodologic issues in child speech disorders research, including measurement methods and classification

criteria for different subtypes of child speech disorders, have been discussed elsewhere (Shriberg, 1993a; Shriberg, 1993b; Shriberg & Austin, 1998; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997a, 1997b). The primary goal of the present study was to estimate the prevalence of speech delay (SD), a classificatory term motivated in the following overview. The secondary goal of the study was to estimate the comorbidity of SD with language impairment, including specific language impairment (SLI).

Nosological Issues

Classification of Child Speech Disorders

Since the paradigm shift in the 1970s from articulatory descriptions to linguistic analyses, a childhood speech disorder of unknown origin has been referenced by many classificatory terms (cf. Edwards, 1997; Elbert, 1997; Shriberg, 1982, 1997). Some of the most common terms are *functional articulation disorder* and *developmental phonological disorder*; hybrids such as *articulation/phonology disorder*; and less theoretically committed terms such as *multiple phoneme disorder*, *speech delay*, or *intelligibility impairment*. The discipline and profession of communicative disorders continues to accommodate this diversity, in large part due to the lack of theoretical clarity on the role of physiological, cognitive-linguistic, and psychosocial deficits as original or maintaining causes of speech disorder. Whichever the preferred classificatory term, some type of developmental speech disorder of unknown origin is universally distinguished from those child speech disorders for which causal origins have been identified.

Typologic and Etiologic Perspectives

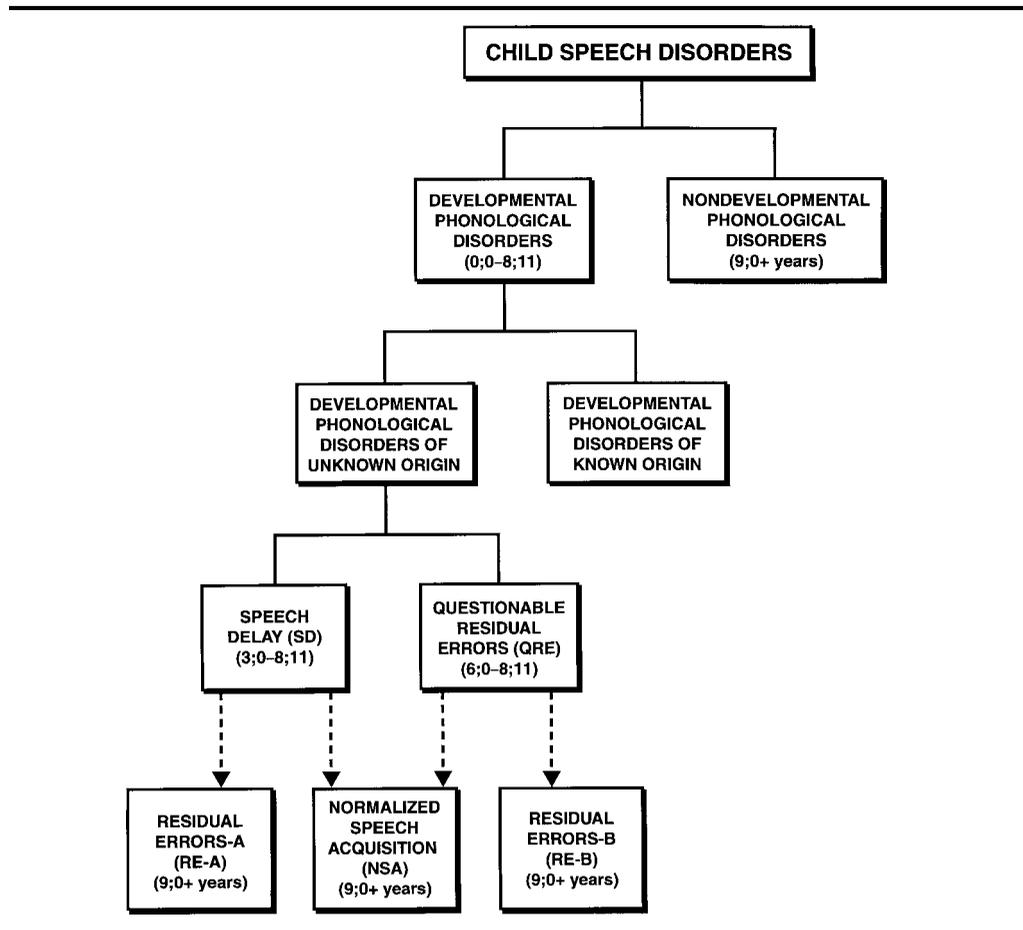
The array of alternative classificatory terms notwithstanding, there is apparent consensus based on typologic speech features for two subtypes of child speech disorders of unknown origin. Figure 1 is an illustration that organizes the primary classification terms proposed in the Speech Disorders Classification System (SDCS; Shriberg et al., 1997b), which is the classification system used for the prevalence study reported in this paper. Rationales for each classification category, including other categories not shown in Figure 1 (e.g., children with suspected apraxia of speech), are provided in Shriberg et al. (1997b). A cover term, *child speech disorders*, is proposed to unify theoretical and applied aspects of speech disorders specific to both developmental and nondevelopmental issues, and to parallel the organization of research and practice in *child language disorders*. The primary division in child speech disorders is between disorders with onsets during the developmental period (developmental phonological disorders, 0;0–8;11 years), which includes phonological development

and fine tuning of articulatory and suprasegmental behaviors, versus those disorders that have their onset after this sociobiological period for the acquisition of speech (nondevelopmental phonological disorders, 9;0+ years). As illustrated in Figure 1, developmental speech disorders are divided into those with unknown versus known origin, the latter including disorders associated with mechanism, cognitive-linguistic, and psychosocial disorders such as the disorders described in the *Diagnostic and Statistical Manual-IV* (American Psychiatric Association, 1994). Children with speech disorders associated with known origins (e.g., craniofacial dysmorphologies, mental retardation, autism) are excluded from prevalence counts of children with speech disorders of unknown origin.

Speech disorders of unknown origin with onsets during the developmental period are divided into the two classifications shown in Figure 1: speech delay (3;0–8;11) and questionable residual errors (6;0–8;11) (cf. Shriberg, 1980; Shriberg & Kwiatkowski, 1980, 1982, 1994). Speech delay (SD) is characterized by age-inappropriate speech-sound deletions and substitutions, typically affecting speech intelligibility. Children with such patterns often have concurrent deficits in language, and some have later deficits in reading and/or spelling (cf. Shriberg & Kwiatkowski, 1994). Questionable residual errors (QRE) is characterized by speech errors limited to clinical distortions (cf. Shriberg, 1993a, Appendix) of one or more fricative, affricate, and/or liquid sounds. Children with this type of speech disorder do not have intelligibility problems, nor are they at higher risk for language deficits (Himmelwright Gross, St. Louis, Ruscello, & Hull, 1985; Prins, 1962; Smit & Bernthal, 1983). This classification category is termed questionable because the relevant literature indicates that such distortion errors are age appropriate during the 3-year period from 6 to 9 years of age.

As illustrated by the classification in the middle of the three boxes at the bottom of Figure 1, some children normalize SD or QRE by 9 years of age (i.e., normalized speech acquisition [NSA]). Because the concept of a speech *delay* is a misnomer after the developmental period, children who have not normalized SD are classified as residual errors-A (RE-A). The term *residual* is used to suggest that any remaining speech errors are residuals of the developmental period. Similarly, children with QRE who retain one or more clinical distortion errors after the developmental period are classified as residual errors-B (RE-B). The dashed lines connecting SD with RE-A, and QRE with RE-B, indicate optional progression to these categories. Notice that the persisting speech errors for RE-A and RE-B could be perceptually identical at this point (i.e., children who were formerly speech delayed could have only distortion errors persisting past 9 years of age). Thus, in a genetics

Figure 1. Overview of relevant categories from the Speech Disorders Classification System (SDCS; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997b).



study assessing the speech of family members of an affected child, historical data might be needed to differentiate speakers with residual errors who previously had SD (i.e., RE-A) from those with former QRE (i.e., RE-B). Studies in progress are attempting to discern if there are acoustic markers of distortions associated with RE-A versus RE-B histories (cf. Flipsen, Shriberg, Weismer, Karlsson, & McSweeney, 1999a, 1999b; Karlsson & Shriberg, 1999).

The hypothesis of two primary forms of child speech disorders of currently unknown origin, SD and QRE, (or, if persistent after 9 years, RE-A and RE-B) is central to the interpretation of prevalence and comorbidity data for speech-genetics studies. As reviewed elsewhere (Shriberg, 1994) and under test in several speech-genetics studies in process, the hypothesis is that only SD is genetically transmitted, whereas QRE (RE-B) arises from environmental variables. If this hypothesis is correct, a near century of prevalence studies that have divided children on the basis of severity (e.g., Hull, Mielke, Timmons, & Willeford, 1971; Wallin, 1916), rather than on the classificatory distinctions reviewed here, do not provide useful data for the needs of speech-genetics studies.

Estimates of the Prevalence of Speech Delay

Critical review of the over 100 relevant prevalence studies of speech disorders worldwide is beyond the scope of the current report. Useful summary data and evaluative reviews include papers by Beitchman, Nair, Clegg, and Patel (1986); Enderby and Philipp (1986); Fein (1983); Hull et al. (1971); Kirkpatrick and Ward (1984); Leske (1981); Morley (1965); Newman (1961); Peckham (1973); St. Louis, Ruscello, and Lundeen (1992); Silva, Justin, McGee, and Williams (1984); and Winitz and Darley (1980). Notwithstanding the considerable effort and careful conduct of many large-scale prevalence studies, especially in the last quarter century, reviews of this prevalence literature have uniformly concluded that there is no consensus on the prevalence of speech disorders in children (cf. Leske, 1981). Methodological critiques include problems associated with population sampling, speech assessment, and, crucially, systems and criteria for classifying children as having a clinically significant speech disorder. Put succinctly over 50 years ago by the authors of a classic descriptive study:

“A great deal of quibbling has gone on as to what constitutes a speech defect” (Roe & Milisen, 1942, p. 37). Four decades and several dozen studies later, Beitchman et al. (1986) observed that “the lack of a comprehensive classification system for speech and language disorders is...a serious barrier to developing useful and accurate prevalence estimates” (p. 98). Since the Beitchman et al. (1986) project, there have been no new large-scale prevalence studies of speech disorders in English-speaking children. As suggested previously, for the present purposes, there has been no published study that has differentiated SD from QRE using the proposed distinctions that define each classification. In those studies in which such distinctions might be inferred or retrieved from the available data, there is no support for the assumption that auditory-perceptual transcription conventions distinguished speech-sound substitutions from speech-sound distortions (e.g., θ/s vs. $[\xi]$), or that systematic empirical criteria were used to differentiate clinical distortions (e.g., $[\xi]$) from nonclinical distortions (e.g., $[s]$) (cf. Shriberg, 1993a, Appendix).

Table 1 includes findings from five of the most widely cited prevalence studies of speech delay of unknown origin in English-speaking children in the age range of present interest. Included are two studies of children in the United States of America and one study each of children in Canada, New Zealand, and the United Kingdom. The studies are arranged by increasing age, from approximately 5 to 8 years. Note the differences in the magnitude of the prevalence estimates, ranging from 2–3% for the 8-year-old children in the “Collaborative Perinatal Study” to 10–13% for the 7-year-old children in “The National Child Development Study.” The source(s) of the discrepancies between these two ranges of estimates can not be readily discerned due to the many

methodological differences among studies reviewed in the previous citations.

Estimates of the Comorbidity of Speech Delay and Specific Language Impairment

Comorbidity refers to “disease(s) that coexist in a study participant in addition to the index condition that is the subject of the study” (Last, 1988, p. 28). It is important to differentiate two sources of comorbidity data in the literature. When computed from epidemiologic surveys, *sample-wide* estimates of the comorbidity of two or more disorders in a population reflect the prevalence of cases in the population sample. When calculated from convenience samples, such as a case records sample or a clinical referral sample, the prevalence of cases comorbid for two or more disorders is the proportion of cases with the *index disorder* who are positive for other disorder(s). The numerators are the same for both types of comorbidity estimates, but the denominators are typically substantially larger for epidemiologic data. Therefore, sample-wide comorbidity is typically lower than when estimated from clinical referral samples of an index disorder. Caron and Rutter (1991) provide a thorough review of the many influences on both types of comorbidity estimates, a major influence being the classification system and criteria used to define the target disorders. In the present context, associations between speech and language disorders have been a fertile area of study (cf. Paul, 1998), including emerging research seeking appropriate phenotypes for speech-language delay. High comorbidity of speech delay and specific language impairment (SLI) would support the likelihood of a common genotype, whereas low comorbidity motivates the need for individual phenotypes for each disorder.

Table 1. Some prevalence estimates for moderate-severe speech delay of unknown origin.

Reference	Source	Country	Age	Prevalence	Increased risk for boys	Sociodemographic factors associated with increased risk
Beitchman, Nair, Clegg, & Patel, 1986	Ottawa-Carleton Study	Canada	5	11%	1.6 ^a	Not reported
Hull, Mielke, Timmons, & Willeford, 1971	The National Speech and Hearing Survey	USA	6	9.7%	1.6	Not reported
Silva, Justin, McGee, & Williams, 1984	Dunedin Multidisciplinary Child Development Study	New Zealand	7	6.5%	2.4	Not reported
Peckham, 1973	The National Child Development Study	England, Scotland, & Wales	7	10–13%	1.5	Socioeconomic class
Winitz & Darley, 1980	Collaborative Perinatal Study	USA	8	2–3%	1.8	Economic; racial/cultural

^aUnadjusted for sensitivity.

There have been few sample-wide, population estimates of the comorbidity of speech and language disorders. The only such estimate available for children of a similar age to those considered in the present study in which somewhat comparable methods were used was reported by Beitchman et al. (1986). In a study of 1,655 Canadian kindergarten children from the Ottawa-Carleton area, Beitchman calculated the sensitivity-adjusted sample-wide comorbidity of speech and language impairment to be 4.6%. For the other type of comorbidity estimates, calculated from samples of children with either SD or SLI, Table 2 provides a summary of findings arranged by increasing age and the index disorder. Methodological information on some of these studies and extended discussion are provided in Shriberg and Austin (1998). For the present purposes, data from the four new studies conducted by and reported in Shriberg and Austin (1998) are limited to those children classified as SD (but only those with *clinical*, rather than *subclinical* disorder; cf. Shriberg & Austin, 1998). Two observations on the comorbidity data in Table 2 are relevant for the present purposes.

First, whether indexed by speech or language disorder, the magnitude of the 24 comorbidity estimates listed in Table 2 varies widely, ranging from 9% to 77%. Second, the magnitude of the estimates does not seem to be strongly associated with either the index disorder or the age of children. Finally, the lowest comorbidity estimates were obtained in the Shriberg and Austin (1998) study (Study 4), which also used a subsample of children with conversational speech data drawn from the present original sample (to be described) but was not referenced to census data on relevant sociodemographic variables. Thus, as concluded in the above discussion of prevalence estimates, subject characteristics and especially methodological characteristics in sampling and classifying speech and language disorder appear to be significant sources of variance in studies of the comorbidity of speech-language disorders. The study reported here attempts to fully address both sampling and measurement needs, toward the goal of obtaining a reliable estimate of the prevalence of speech delay and its comorbidity with language impairment.

Table 2. Comorbidity estimates for speech-language disorders.

Study	n	Comorbidity estimates by index disorder		
		Mean age ^a	Speech index	Language index
Paul (1993)	37	3.0 ^b		37%
Whitehurst et al. (1991)	22	3		35%
Connell, Elbert, & Dinnsen (1991)	23	4	43%	
Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber (1986)	33	4	60%	
Shriberg & Kwiatkowski (1994)	64	4	66%	
Bishop & Edmundson (1987)	66	4.0 ^c		74
Paul (1993)	37	4.0 ^b		16%
Shriberg & Austin (1998, Study 1)	44	4	41%	
Shriberg & Austin (1998, Study 2)	35	4	11%	
Tallal, Ross, & Curtiss (1989)	76	4		60
Bishop & Edmundson (1987)	67	4.5 ^c		55
Shriberg et al. (1986)	38	5	50%	
Bishop & Edmundson (1987)	68	5.0 ^c		34
Shriberg & Austin (1998, Study 3)	23	5	39%	
Tomblin (1996b)	862	5		25
Paul & Shriberg (1982)	30	6	66%	
Shriberg & Austin (1998, Study 4)	79	6		9%
Shriberg & Kwiatkowski (1982)	43	6	77%	
Schery (1985)	718	7	75% ^d	75% ^d
St. Louis, Ruscello, Grafton, & Hershman (1994)	20	7	45%	
St. Louis et al. (1994)	20	7		65%
Ruscello, St. Louis, & Mason (1991)	24	12.5	54% ^f	
Ruscello et al. (1991)	24	12.5	21% ^g	

^aAges rounded up at 0.5, unless the mean age was provided within a narrow age range. ^{b,c}Same children followed longitudinally. ^dNo index disorder in this survey. ^eEstimated from grade level. ^fClassified as having delayed articulation. ^gClassified as having residual errors.

Methods

The Epidemiology of Specific Language Impairment (EPISLI) Project

Participants in the present study were a subsample of the 7,218 children whose language and speech were assessed in the National Institute on Deafness and Other Communication Disorders contract titled "Epidemiology of Specific Language Impairment" (EPISLI; Tomblin et al., 1997; Tomblin, Records, & Zhang, 1996). The primary goal of the EPISLI grant was to estimate the prevalence of SLI in the United States of America. Readers are encouraged to consult the two references cited for detailed descriptions of methods and findings. The following is a brief overview of the project.

Participants

The EPISLI research project was designed to assess the language status of monolingual, English-speaking kindergarten children living in three population centers in Iowa and Illinois: Des Moines, Cedar Rapids/Waterloo/Cedar Falls, and the "Quad Cities" that straddle the Mississippi River (Davenport, IA; Bettendorf, IA; Moline, IL; and Rock Island, IL). The three centers each include urban, suburban, and rural residential strata, as determined from 1990 census data on population densities and distances of residences from the urban center of each city. Participants whose speech and language were screened using procedures described below included 7,218 children in 21 public school districts, with 655 to 1,001 children in each of the nine clusters defined by the three geographic sites and the three residential strata within each site.

Method

The EPISLI cross-sectional project included two phases in which eligible children were screened with a brief language screening test, and subsamples of those children were subsequently administered a diagnostic battery. Of an original 7,844 children, 7,218 met eligibility criteria and were available for screening. Children in kindergarten classrooms were screened using 40 items from the Picture Vocabulary, Sentence Imitation, and Grammatical Completion subtests of the Test of Language Development-2: Primary (TOLD-2: P; Newcomer & Hammill, 1988). Scoring guidelines were modified to be appropriate for speakers of African American English.

All children who failed the screening, plus a random sample of 33% of children who passed the screening were

invited to continue in the diagnostic testing phase if they came from monolingual English home backgrounds and did not have histories indicating one or more of an array of disabilities. This approach would maintain in the subsample the same ratio of children passing (74.7%) and failing (25.3%) the language screening. Of the 3,877 children who were eligible for the diagnostic phase, 2,084 were granted parental consent to participate in the diagnostic assessment. Analyses of potential bias associated with participation indicated no significant associations between participation/nonparticipation and relevant study variables (cf. Tomblin et al., 1997). The diagnostic battery, which included measures of hearing, impedance, intelligence, language, and speech, was administered by examiners in specially equipped vans parked on each site. As described in Tomblin et al. (1996), classification as SLI required scores at the 10th percentile or lower on two or more of five composite scores developed from seven language measures. Information on examiner training and supporting examiner reliability is provided in the two EPISLI references; Tomblin et al. (1997) includes all findings relevant to the prevalence of SLI. As extrapolated from the EPISLI project findings, 7.4% of American kindergarten children were estimated to have SLI.

The Prevalence of Speech Delay Study

Participants

To pursue associated questions in the EPISLI project addressing sociodemographic and risk issues, Tomblin and colleagues assembled a subsample of the 2,084 children who were tested with the diagnostic battery. A subsample of 1,456 children (20.2% of the original 7,218 screened children; 69.9% of the 2,084 children receiving the diagnostic battery) was constructed to be representative of the original 1990 census data on the following variables: *sex* (census: 49% female, 51% male; subsample: 44% female, 56% male); *strata* (census: 83.5% urban, 16.5% rural; subsample: 85.5% urban, 15.5% rural); *race* (census: 87% white, 13% black; subsample: 88.4% white, 11.6% black); *education level* estimated by residential blocks (census: 3.6 [high school degree = 4], subsample: 3.9); *income per capita* (census: \$14,420, subsample: \$13,150); and *ratio of income to poverty threshold* (census: 1.74, subsample: 1.71). Additional comparisons indicated that the subsample of 1,456 children was comparable to the sample of 7,218 children on the *language screening pass/fail rate* (sample: 74.6% pass, 25.4% fail; subsample: 74.7% pass, 25.3% fail), *language screening scores* (sample: mean = 19.9; standard deviation = 7.88; subsample: mean = 20.1; standard deviation = 7.68), and *strata* divided into three levels (sample: 38% urban, 30% suburban, 32% rural;

subsample: 41% urban, 29% suburban, 30% rural).

Although not used to estimate the prevalence of SLI in the EPISLI project, the Word Articulation subtest of the Test of Language Development–2: Primary (TOLD–2: P; Newcomer & Hammill, 1988) was included in the diagnostic assessment protocol administered to the 2,084 children. Examiners were instructed to record on the Word Articulation subtest form the type of articulation error for the target sounds in each of the 20 words. For 127 children, TOLD information was incomplete or otherwise unusable for the needs of the present study, thereby reducing the original sample size from 1,456 to 1,328 (91%) children.

Children’s scores on the 20-item Word Articulation subtest were found to be insufficient as the sole source for a valid estimate of the prevalence of SD in this sample of 6-year-old children. Anticipating this constraint on the potential use of EPISLI data to estimate the prevalence of SD, the authors arranged to obtain approximately 500 audiocassette-recorded conversational

speech samples from children participating in pilot studies and in the final EPISLI project. Conversational speech samples were obtained for 303 of the 1,328 children in the sociodemographically representative subsample. Thus, there were 303 children for whom both TOLD and conversational speech samples were available, and an additional 1,025 children with only TOLD data. However, these 303 children were not themselves representatively distributed across the sociodemographic variables on which the sample of 1,456 children had been assembled.

Table 3 is a description of the 1,328 subsample participants by sex, residential strata, and racial/cultural background. As shown in the bottom line of Table 3, the sample included 11% more boys (55.5%) than girls (44.5%). As indicated in the subtotals for each stratum, there were approximately equal proportions of rural (29.1%) and suburban (28.2%) children, but nearly 15% more children from the urban stratum (42.7%). These differences among sex and strata were needed to match

Table 3. Description of the 1,328 participants by sex, residential strata, and racial/cultural background.

Strata	Racial/cultural background	Boys		Girls		All	
		<i>n</i>	% ^a	<i>n</i>	%	<i>n</i>	%
Rural							
	American Indian	0	–	1	0.1	1	0.1
	Asian	1	0.1	0	–	1	0.1
	African American	3	0.2	3	0.2	6	0.5
	Hispanic	2	0.2	2	0.2	4	0.3
	White	226	17.0	149	11.2	375	28.2
	All	232	17.5	155	11.7	387	29.1
Suburban							
	American Indian	0	–	1	0.1	1	0.1
	Asian	3	0.2	1	0.1	4	0.3
	African American	20	1.5	23	1.7	43	3.2
	Hispanic	2	0.2	1	0.1	3	0.2
	White	200	15.1	123	9.3	323	24.3
	All	225	17.0	149	11.2	374	28.2
Urban							
	American Indian	1	0.1	3	0.2	4	0.3
	Asian	0	–	5	0.4	5	0.4
	African American	43	3.2	58	4.4	101	7.6
	Hispanic	5	0.4	2	0.2	7	0.5
	White	230	17.3	219	16.5	449	33.8
	All ^b	280	21.0	287	21.6	567	42.7
All							
	American Indian	1	0.1	5	0.4	6	0.5
	Asian	4	0.3	6	0.5	10	0.8
	African American	66	5.0	84	6.3	150	11.3
	Hispanic	9	0.7	5	0.4	14	1.1
	White	656	49.4	491	37.0	1147	86.4
	All	737	55.5	591	44.5	1328	100.0

^aPercentage of the data set. ^bRace for one male urban child did not clearly fit any of the five classifications.

the nationwide racial/cultural background compositions. As pointed out by Tomblin and colleagues (1997), the EPISLI samples deviated from the national population characteristics in that they had fewer Asians, Hispanics, and Native Americans. This deviation was likely associated with the greater rates of languages other than English being spoken in these homes, thereby eliminating children from participation due to the monolingual English inclusionary criterion.

Classification of Children as Speech Delayed (SD) From Conversational Speech Samples

Speech Sampling

The EPISLI examiners were trained to obtain a conversational speech sample following a protocol that has been used in prior studies of children with typically developing and delayed speech (e.g., Shriberg & Kwiatkowski, 1994). The protocol included procedures to obtain technically adequate audiocassette recordings and to evoke rich speech samples from the children. Sony 5000 audiocassette tape recorders (calibrated to specifications) with matching University Sound 659L external microphones were used to obtain all speech samples. Several rounds of pilot testing were conducted so that feedback could be provided to all examiners on specific procedures to maximize signal quality and linguistic productivity for the planned speech analyses. Subsequent analyses of the 303 samples used in the present study indicated a mean of 235 usable words per sample ($SD = 45$, range = 89–499), which met sufficiency guidelines for all speech analyses.

Phonetic Transcription

High-quality copies of the original audiocassette tapes were forwarded to the Madison site for transcription and computer processing. The transcription protocol included guidelines to maximize the reliability of auditory-perceptual data, including phoneme- and allophone-level symbolization rules relevant to classification of SD (Shriberg, 1986). Two research assistants transcribed the audiocassette tapes over a 2-year period.

Transcription Agreement

Ten conversational samples were randomly selected from a group of 92 speech samples (11%) to assess transcription agreement. The 92 tapes represented the point at which approximately 20% of an eventual 510 tapes were transcribed. Interjudge and intrajudge transcription reliability was calculated by a computer program that computes broad (i.e., disregarding diacritics used by either transcriber) and narrow (including diacritics used by either transcriber) point-to-point transcription

agreement for consonants and vowels/diphthongs (Shriberg & Olson, 1988).

Findings for interjudge agreement were as follows: Broad transcription agreement was 90.9% for consonants and 87.5% for vowels/diphthongs, and narrow transcription agreement was 82.5% for consonants and 80.5% for vowels/diphthongs. Findings for each of the two transcriber's intrajudge agreements were as follows: Broad transcription agreement was 93.7% and 91.6% for consonants, and 92.4% and 90.4% for vowels/diphthongs; narrow transcription agreement was 86.6% and 85.3% for consonants, and 87.2% and 85.6% for vowels/diphthongs. These transcription reliability estimates are comparable to those reported in other studies in normal and disordered child phonology (McSweeney & Shriberg, 1995; Shriberg & Lof, 1991), and were considered adequate for the purposes of the present study. As described below, the computer program that classifies children's speech status uses information from both consonant and vowel/diphthong sound classes, as well as diacritic-level sound changes.

Classification of Children as Speech Delayed (SD) From TOLD Scores Validated to Conversational Speech

The Speech Disorders Classification System (SDCS) described previously was used to classify the speech status of the 303 children based on their error patterns in the speech sample. Using the reference database described in Shriberg (1993a, Appendix), the computer program determines if the child's speech pattern meets criteria for several speech classifications (cf. Shriberg et al., 1997b). Because 6 years is the earliest age at which children are eligible to be classified as having questionable residual errors (QRE; e.g., /s/ and /r/ errors are not age-inappropriate until age 6), the validity of QRE classifications was a concern with the present sample. Many of the children had not quite reached 6 years of age, and most were just a few months older. Thus, the probability of both false positives and false negatives was considered high for all QRE classifications.

A series of feasibility studies was conducted to determine if a procedure could be developed to use information on the TOLD score sheets to classify children as meeting SDCS criteria for SD. Because both TOLD and conversational speech data were available for 303 participants, the goal was to develop an algorithm with sufficient sensitivity and specificity to provide valid speech classifications for the remaining 1,025 children in the EPISLI subsample. The Appendix summarizes methods and findings in each of the three phases of this study series, including two pilot studies testing possible algorithms, a criterion-validation analysis to assess sensitivity and specificity for the most promising procedure,

and the procedures developed for the prevalence study. Due to the validity concerns for QRE at 6 years of age described above, as well as the lack of available reliability data for the EPISLI examiners' transcription of distortions on the TOLD forms, the algorithms focused only on discriminating SD from NSA. Accordingly, no prevalence data will be reported or were assembled for QRE. Two research assistants using the finalized coding protocol determined if each of the 1,025 children's TOLD responses met criteria for SD.

Results

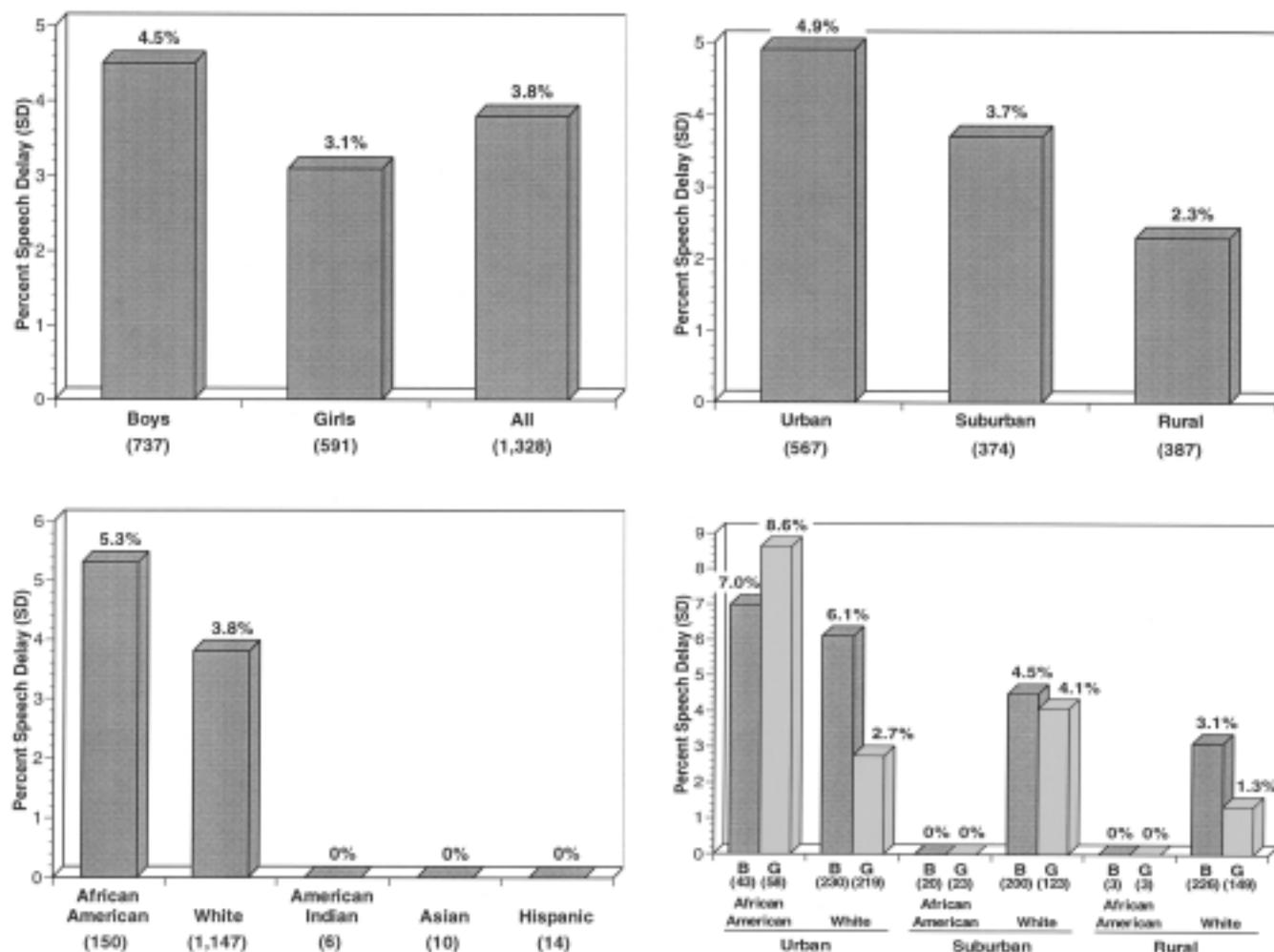
Prevalence of Speech Delay

Figure 2 is a summary of the prevalence findings separately by sex (top left panel), residential strata (top right panel), racial/cultural background (bottom left panel), and cross-tabulations among these three variables (bottom right panel).

The number of children for each estimate is provided below each bar. Beginning with the top left panel, the overall prevalence of SD in this sample of 6-year-old children was 3.8% (95% confidence intervals [Fleiss, 1981] in parentheses here and hereafter: 2.9–5.0%). A total of 4.5% (3.2–6.3%) of the boys and 3.1% (1.9–4.9%) of girls met criteria for SD, yielding a boys-to-girls ratio of 1.5:1. Prevalence rates for SD by residential strata were 4.9% (3.3–7.1%) for urban, 3.7% (2.1–6.3%) for suburban, and 2.3% (1.1–4.5%) for rural strata. Prevalence rates for SD by racial/cultural background were 5.3% (2.5–10.6%) and 3.8% (2.8–5.1%) for children identified as African American and White, respectively, and 0% for the very few children identified as American Indian, Asian, or Hispanic.

The bottom right panel in Figure 2 includes 12 prevalence estimates reflecting cross tabulations among residential strata, racial/cultural background, and sex.

Figure 2. Prevalence of speech delay (SD) in a sample of 1,328 six-year-old children. The four panels provide information for the variables of sex (top left), residential strata (top right), racial/cultural background (bottom left), and for the three-way associations among variables (bottom right).



Small cell sizes for most estimates limit the utility of confidence intervals around means and constrain all inferences about the relative strength of each of the three variables. The following three observations of possible trends warrant brief comment for future study.

First, in contrast to the unconditional 1.5:1 boys-to-girls prevalence ratio for SD shown in the upper left panel, sex ratios in the lower right panel range from 1:1.2 (urban African American children) to 2.3:1 (urban White children), and 2.4:1 (rural White children) in those comparisons where at least some children of either sex were affected. Second, although urban children of either sex generally had higher prevalence of SD compared to children from suburban and rural residential strata, there was one exception (urban White girls = 2.7%, suburban White girls = 4.1%). Finally, these prevalence estimates also suggest possible interdependencies between racial/cultural background and residential strata. As shown, African American boys and girls from urban backgrounds had the highest estimated prevalence of SD among all groups. There were no African American children from suburban ($n = 43$) or rural ($n = 6$) backgrounds identified as SD. Again, the generalizability of these data needs to be assessed in studies using appropriately large study samples.

Comorbidity of Speech Delay and Language Impairment

Sample-Wide Comorbidities

Table 4 provides data on the prevalence of language impairment in this sample of 737 boys and 591 girls (top rows) and comorbidities of speech delay with normal cognition/language impairment (i.e., specific language impairment), low cognition/language impairment, and

low cognition/normal language (bottom rows). Tomblin et al. (1997) describe the procedures used to assess and classify cognitive status and language impairment. All children with normal cognition were required to have a performance IQ greater than 87, as calculated from the Block Design and Picture Completion subtests of the Wechsler Preschool and Primary Scale of Intelligence-Revised (Wechsler, 1989). Children termed low cognition had IQs between 70 and 87. As with the speech data, classification of children as language impaired is sensitive only to those children who have not normalized by this age (i.e., the data in this study do not reflect the incidence of speech delay or language impairment). The prevalence data for language impairment for the present subsample of 1,328 children (Table 4, top rows) differ slightly from the prevalence estimates for the complete sample reported in Tomblin et al. (1997).

The sample-wide comorbidity estimates in the right half of Table 4 include percentages for each sex and averaged over sex (to adjust for unequal numbers of boys and girls). The comorbidity of speech delay with each of the three cognitive-linguistic classifications was less than 1% (CIs range from .05–1.9%). Summing over the two classifications of language impairment, specific language impairment and low cognition/language impairment, 1.3% (.79–2.1%) of these children remained comorbid for speech delay and language impairment at 6 years of age.

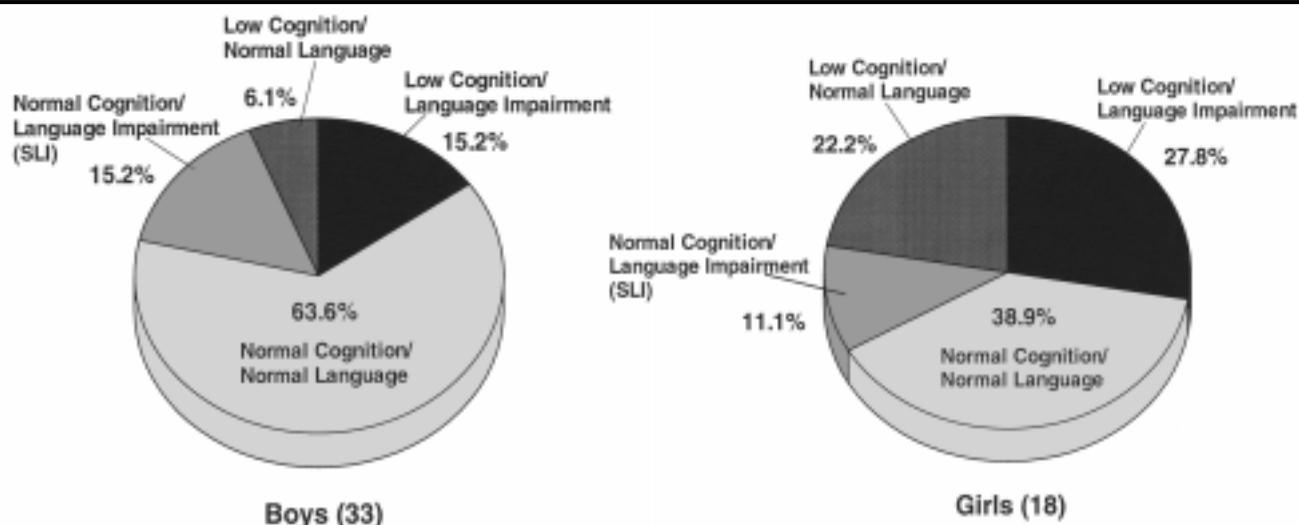
Comorbidity Indexed on Speech Delay

Figure 3 is a summary of the comorbidity data as indexed by the 51 children identified as SD. For sample sizes this small (boys, $n = 33$; girls, $n = 18$), the 95% confidence intervals around these comorbidity estimates were

Table 4. Prevalence of language impairment in the present sample and sample-wide comorbidity of speech delay with normal cognition/language impairment (i.e., specific language impairment), low cognition/language impairment, and low cognition/normal language. 95% confidence intervals for each estimate are given in parentheses.

	<i>n</i>	Prevalence						Comorbidity/cooccurrence					
		Normal cognition/ language impairment (SLI)		Low cognition/ language impairment		Low cognition/ normal language		Speech delay + normal cognition/language impairment (SLI)		Speech delay + low cognition/ language impairment		Speech delay + low cognition/ normal language	
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Boys	737	66	9.0 (7.1–11.4)	38	5.2 (3.8–7.1)	80	10.9 (8.8–13.4)	5	0.68 (.25–1.7)	5	0.68 (.25–1.7)	2	0.3 (.05–1.1)
Girls	591	42	7.2 (5.3–9.7)	28	4.7 (3.2–6.8)	66	11.2 (8.8–14.1)	2	0.34 (.06–1.4)	5	0.85 (.13–1.6)	4	0.7 (.23–1.9)
Average			8.1 (6.7–9.7)		5.0 (3.9–6.3)		11.9 (9.5–12.9)		0.51 (.22–1.1)		0.77 (.39–1.5)		0.5 (.21–1.1)

Figure 3. Comorbidity of speech delay (SD) as the index disorder with three categories of cognitive-linguistic impairment: normal cognition/language impairment (i.e., specific language impairment), low cognition/normal language, and low cognition/language impairment.



large. The percentages of boys and girls with concurrent SD and SLI (i.e., normal cognition/language impairment) were essentially similar (15.2% [5.7–32.7%] and 11.1% [3.4–27.9%], respectively). However, the percentages of boys and girls with normal cognitive status were notably different. Among the 33 boys with SD, over 60% (63.6% [45.1–79.0%]) had normal cognition/normal language, whereas approximately 40% of the 18 girls (38.9% [18.3–63.9%]) had normal cognition/normal language at this age. The approximately 20% difference was reflected in the remaining two classifications. Among the boys with SD, 15.2% (5.7–32.7%) had low cognition/language impairment, and an additional 6.1% (1.1–21.7%) had low cognition/normal language, totaling 21.3% (9.7–39.4%) with low cognitive status. For girls with SD, 27.8% (10.7–53.6%) had low cognition/language impairment and an additional 22.2% (7.4–48.1%) had low cognition/normal language. Thus, a total of 50% (26.8–73.2%) of the girls with SD were found to have low cognitive status, over twice the percentage of the boys with SD with low cognitive status (21.3% [9.7–39.5%]).

Comorbidity Indexed on Cognitive-Linguistic Status

Figure 4 is a summary of the comorbidity data using cognitive-linguistic status as the index disorder, which requires individual percentages for each of the four types of cognitive-linguistic status. As shown in the upper left panel in Figure 4, for children with Normal Cognition/Normal Language at kindergarten age, SD was identified in 3.8% (2.4–5.8%) of the boys and in 1.5% (.65–3.2%) of the girls. As shown in the upper right panel, which provides data for children with SLI (i.e., normal cognition/language impairment), SD was identified in

7.6% (2.8–17.5%) of the boys and 4.8% (8.4–17.5%) of the girls. For children with low cognition/normal language (lower left panel), SD was identified in 2.5% (4.3–9.6%) of the boys and 6.1% (2.0–15.6%) of the girls. Finally, for children with low cognition/language impairment (lower right panel), SD was identified in 13.7% (5.3–29.5%) of the boys and in 17.9% (6.8–37.6%) of the girls.

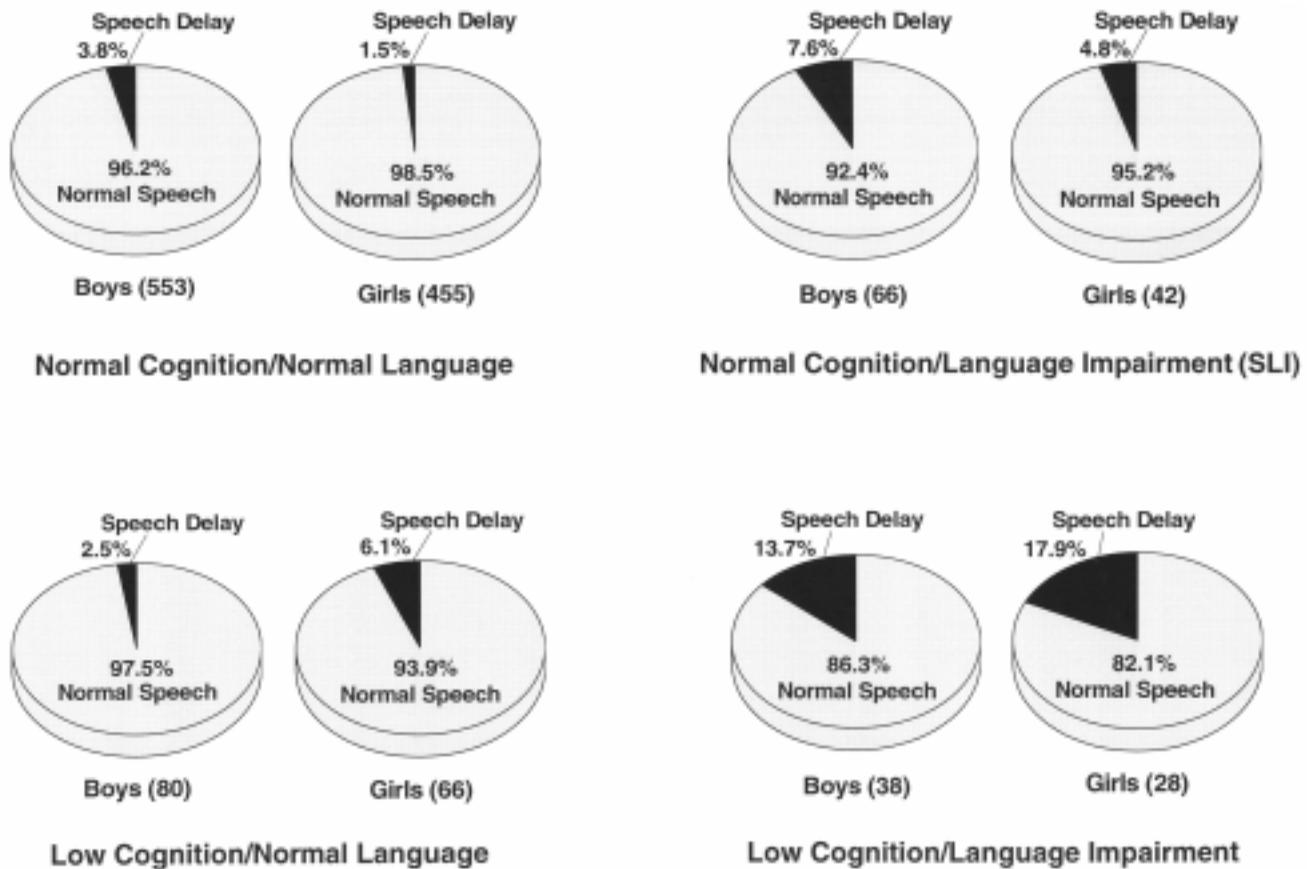
Discussion

Concurrent Validity Support for the Obtained Prevalence Estimate

Before proceeding to substantive issues raised by these data, it is useful to review several sources of validity support for the 3.8% prevalence estimate for speech delay in 6-year-old children obtained in this study. Compared to extrapolations from Leske's (1981) data suggesting 2.5–3% prevalence of speech-language disorders in 3- to 5-year-old children, the obtained estimate for 6-year-old children may be viewed as unexpectedly high. It is also higher than expected based on the rates for language delay in preschool age (i.e., 3–8% at age 3; Silva, 1980; Stevenson & Richman, 1976). However, compared to Hull et al. (1971) or Beitchman et al. (1986), who reported 9–11% prevalence rates for first-grade children, the 3.8% estimate might be considered low. Finally, the 3.8% estimate obtained in the present study is generally consistent with the 2–3% prevalence estimates for children of this age extrapolated from reports by Winitz and Darley (1980) and Enderby and Philipp (1986).

As reviewed previously, the primary sources of variance in prevalence estimates include the epidemiologic

Figure 4. Comorbidity of three categories of cognitive-linguistic impairment as the index disorder with speech delay (SD): normal cognition/language impairment (i.e., specific language impairment), low cognition/normal language, and low cognition/language impairment.



methods used to select population samples and the diagnostic criteria used to identify the target disorder. On the first set of considerations, Tomblin et al. (1997) provide extended comparative commentaries on the cluster sampling techniques used in the EPISLI study in relation to those used in other epidemiologic studies of speech-language disorders. Although cell size limitations constrain generalizations of these prevalence figures at the level of specific racial/cultural groupings, the U.S. population sampling methods are viewed as generally robust. On issues associated with the diagnostic criteria used to identify children with SD, we point to several sources of information both internal and external to this study that provide criterion-validity support for a 3.8% prevalence estimate for SD.

Internal Support

Internal validity support for the claim that children classified as SD had significant speech involvement can be examined by inspecting the Percentage of Consonant Correct-Revised (PCC-R; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997a) data for 16 of the 51 children with speech delay for whom conversational

samples were available for the prevalence study compared to the remaining 287 (of the 303) children classified as having normal speech acquisition (NSA). A PCC-R score is the percentage of correct consonants in a conversational speech sample, with only omissions and substitutions counted as errors (i.e., distortion errors are scored as correct; cf. Shriberg et al., 1997a). Thus, unlike the Percentage of Consonants Correct (PCC) metric, which scores speech-sound distortions as incorrect, the PCC-R reflects severity differences based only on those cognitive-linguistic errors that define SD. Thus, in scoring distortions as correct, the PCC-R does not discriminate NSA from QRE, the two classifications that were merged for the purposes of the present study. The mean PCC-R score for the 16 children classified as SD on the SDCS was 78.7 (standard deviation = 7.77), compared to a PCC-R mean of 92.9 (standard deviation = 4.45) for the 287 children classified as NSA. A *t* test of the difference in PCC-R scores was statistically significant ($t = 7.30, df = 15, p < .001$). Thus, as sampled from their conversational speech, the 6-year-old children classified as SD in this subsample averaged approximately 15 percentage points (95% confidence interval = 10.13–18.50 percentage points) lower than the

children classified as having typically developing speech.

External Support

External support for the validity of the 3.8% prevalence estimate can be marshaled from convergent information from several sources. Campbell (personal communication, 1998) has used the SDCS to classify the speech status of 241 3-year-old children. Campbell's demographically diverse population sample in the Pittsburgh area was ascertained from clinics and pediatric practices in the context of a large-scale longitudinal study. With 34 of the 241 children's conversational speech samples classified as SD by the SDCS program, the prevalence of SD in this culturally diverse cohort of 3-year-old children was 14.1%. This estimate for 3-year-old children could be used to provide a prevalence estimate for 6-year-old children if an estimate were available on the percentage of children with SD at 3 years of age who normalize SD by 6 years of age. Such estimates are available from two sources.

Based on clinical observations of several hundred children treated at a university speech clinic, Shriberg (1994) estimated that approximately 75% of preschool children with SD normalize their speech by 6 years of age. This retrospectively based normalization estimate has recently gained support in preliminary findings from a prospective study of speech-sound normalization in 20 children identified at 3.5–4.5 years of age as SD using the SDCS program (Shriberg, McSweeney, & Kwiatkowski, 1999). Based on the 17 children for whom sufficient longitudinal data are available, 12 (70.6%) of the speech-delayed children had normalized speech by 6 years of age (plus or minus 2 months). Thus, when the estimate for normalization by age 6 based on the retrospective data (75%) is applied to Campbell's 14.1% prevalence estimate for SD in his 3-year-old children, the prediction for 6-year-old children is in close agreement with the present prevalence estimate for 6-year-olds. Specifically, the remaining 25%, or a total of 3.5% of Campbell's children ($.25 \times .141 = .035$), would be expected to retain SD by 6 years of age, which is within three tenths of the 3.8% prevalence figure for SD obtained in the present study. An estimate based on the prospective study data (70.6%) yields a 4.2% prevalence of SD in Campbell's children at 6 years, within four tenths of the 3.8% prevalence figure for SD obtained in the present study.

Finally, an additional source of external validity support for the 3.8% figure is available in an analysis of the Irwin, Huskey, Knight, and Oltman (1974) data, reported by Beitchman et al. (1986). Using recovery by third grade as the criterion to identify children with prior speech delay (including children whose only remaining errors were speech-sound distortions), which occurred

for 82% of children from first to third grade in the Irwin et al. (1974) study, Beitchman and colleagues estimated the prevalence of moderate to severe phonological disorders in first grade children to be 3.8%. Thus, although the present prevalence estimate for children of this age is considerably less than the incidence of speech delay as assessed at earlier ages (with incidence the central datum for genotype issues), the procedures used for phenotypic classification appear to have substantial empirical support.

Prevalence

The present finding of 3.8% speech delay in 6-year-old children appears to be the first prevalence estimate validated to the classification criteria described previously (i.e., persisting deletion/substitution errors) as assessed in conversational speech. Assuming validity support for these methods (see above) and that 3.8% is a conservative prevalence estimate (i.e., unadjusted for false negatives; see Appendix), a significant percentage of children begin early elementary grades with two potentially handicapping conditions. From a phonological perspective, persistent deletion/substitution errors in children of this age have been associated with an increased probability of problems in reading, spelling, and other academic areas, perhaps throughout the life span (cf. Felsenfeld, McGue, & Broen, 1995). From a psychosocial perspective, persistent speech-production problems have also been associated with increased risk for social acceptability in school and on the playground (e.g., Crowe Hall, 1991; Silverman & Paulus, 1989). As indicated by the findings in Figure 2 and considered below, these two potential correlates and consequences of speech delay may be significantly moderated by several subject factors.

Sex

The boys-to-girls ratio of speech delay found in the present study is greater than 1 (1.5:1), but lower by as much as 100% than estimates ranging from 2:1 to 3:1 in samples of preschool children with speech delay (cf. Shriberg & Kwiatkowski, 1994). Due to the variations of values in the research literature, it is not clear whether the present lower boys-to-girls prevalence ratio is due to the identification of relatively fewer boys with SD in this population sample or to the identification of relatively more girls with SD.

One methodological possibility consistent with a lower percentage of identified boys with SD concerns the SDCS inclusionary requirement of age-inappropriate deletion/substitution errors, with distortion errors not included in this criterion. A speculation in the speech literature is that boys have more distortion errors than

girls because boys lag girls in motor speech development, and speech-sound distortions reflect delays or deficits in motor speech processes (e.g., Silva et al., 1984). If at least the first of these premises is true, fewer boys might have been identified in the present compared to past studies because distortion errors do not qualify for SD in the SDCS approach. Close analysis of children with SD in our database does not support the hypothesis of more or relatively more (i.e., corrected for overall severity level) speech-sound distortions in boys compared to girls, although there are some trends for distortions of individual speech sounds. Moreover, Silva and colleagues' (1984) analysis of the gross and fine motor performance of children with speech-language disorders and normal cognitive status compared to children with typically developing speech failed to find statistically significant between-group differences (however, see Goffman, 1998; Stark & Tallal, 1988).

A second possible explanation for the lowered boys-to-girls ratio for speech delay in the present study compared to past estimates is that this estimate does not have the potential for ascertainment bias that occurs in clinical referral samples (e.g., Shaywitz, Shaywitz, Fletcher, & Escobar, 1990). The ascertainment bias hypothesis is that, due to differences in perceptions and expectations of behavioral issues in boys and girls, boys are overidentified in clinical referral samples, and girls are underidentified. Tomblin et al. (1997) discussed this possible explanation for the 1.3:1 boys-to-girls ratio found for SLI in the original sample for the present study, which is also lower than values obtained in other prevalence studies of language impairment and is close to the 1.5:1 ratio obtained for speech delay in the present study. All of the sex ratios of 2.5:1 or greater in studies of speech delay summarized in Shriberg and Kwiatkowski (1994) were obtained from clinical referral samples. Also on this point, research using clinical referral samples typically has exclusionary criteria to rule out children with cognitive deficits. In the present study, 50% of girls with SD had low cognitive status, compared to 21.3% of the boys. Thus, the lowered boys-to-girls prevalence ratio could, in part, reflect differences in the cognitive status of children included in other studies.

A third perspective on the present finding is the possibility that boys and girls may have different rates and time courses of normalization of speech delay. Higher boys-to-girls ratios reported in prior studies have typically been obtained from preschool samples, generally weighted more by children in the 4- to 5-year age range (Shriberg & Kwiatkowski, 1994). There are few available data on the rates of normalization of speech delay in boys versus girls, with or without treatment. As indicated previously, retrospective studies (Shriberg, Gruber, & Kwiatkowski, 1994; Shriberg, Kwiatkowski, & Gruber, 1994) and a prospective study in process suggest that

approximately 75% of children with speech delay normalize by 6 years of age, the age at which prevalence was estimated in the present study. Also, on this issue, the stability of diagnosis of speech and language delay at different ages is an important measurement consideration in prevalence studies. At least a portion of the variance in boys-to-girls prevalence ratios at different ages could be due to unstable psychometric classifications across normative reference data. Finally, as addressed in the next section, sex ratios may be further complicated by interactions with sociodemographic factors.

Sociodemographics

The sociodemographic data in the top right and bottom two panels of Figure 2 suggest that sociodemographic variables are strongly associated with the prevalence of speech delay, especially the data on strata (urban, suburban, rural). Such associations are well documented in the child speech literature. In a comprehensive review of the classical research on articulation disorder and perceptual, articulatory, cognitive, and social variables, Winitz (1969) concluded that social variables accounted for the greatest variability in speech-sound acquisition. In a review of findings from the "Collaborative Perinatal Study" (Lassman, Fisch, Vetter, & LaBenz, 1980)—the largest American study of risk factors for communicative disorders considering the number of sites, children, and variables studied—Winitz and Darley (1980) also concluded that socioeconomic variables (and IQ at 4 years of age) contributed most to the variance in articulation and intelligibility at 8 years of age. As suggested by the estimates in Figure 2, this figure may be twice as high or higher in populations of children from diverse backgrounds associated with socioeconomic disadvantage.

One possible methodological explanation for the range of prevalence estimates associated with sociodemographic variables is that it might reflect measurement bias at one or more stages of data collection and data reduction. All the examiners in the EPISLI project were White females. Although guidelines were developed to account for sociodemographic differences at the data reduction phase (cf. Tomblin et al., 1997), the response to testing by children who were from different racial and cultural backgrounds than those of the examiners may have been biased toward less proficiency than if sampled using other methods and obtained by examiners from similar backgrounds.

Although test bias cannot be excluded as a possible explanation for speech and language differences associated with racial/cultural backgrounds, it cannot explain the relatively large prevalence differences across residential strata for children with similar racial/cultural

backgrounds. Tomblin (1996b) in an overview of sociodemographic findings for children with SLI, concluded that parental education and income were the significant mediating variables in these data. Winitz and Darley (1980) also reported that mother's education was significantly associated with speech development. Subanalyses of mother's education and speech delay could not be pursued in the present study because a disproportionate percentage of participants from lower socioeconomic levels had elected not to participate in the extensive telephone interview from which these data were obtained.

Comorbidity

The sample-wide comorbidity data in Table 4, which indicate a less than 2% (CIs range from .21–1.5%) co-occurrence of speech and language impairment for the 6-year-old children in this sample, are consistent with or marginally higher than the comorbidity percentage expected by chance. For example, the chance probability for the occurrence of SD and SLI in this sample is .31% (i.e., prevalence of SD (3.8%) \times prevalence of SLI (8.1%) = .31%; Caron & Rutter, 1991), whereas the obtained comorbidity was .51% (Table 4). As reviewed previously, the only available sample-wide comorbidity estimate for children close in age (5 years) to participants in the present study (6 years) in which somewhat comparable methods were used were the data reported by Beitchman et al. (1986). Beitchman calculated the sensitivity-adjusted, sample-wide comorbidity of speech and language impairment in 1,655 Canadian children from the Ottawa-Carleton area to be 4.6%.

The absolute magnitudes of each comorbidity estimate, indexed by speech delay or language impairment, are also lower than reported previously for children in this age range. For the 10 studies providing comorbidity data for children ranging in age from 5–7 years in Table 2, estimates ranged from 9% to 77%. Nine of the 10 estimates were well above the 5–15% comorbidity values obtained in the present study (see Figures 3 and 4). For comparison, Shriberg and Austin (1998) reported comorbidity for several preschool-age samples ranging from 40–60%. Two possible explanatory factors warrant comment.

First, comorbidity rates of both types—sample-wide comorbidity and by index disorder—would obviously have been higher if the obtained prevalence rate for SD were higher. As described in the Appendix, a conservative estimate of the prevalence rate for speech delay was determined to be preferable to a statistically adjusted estimate. The unadjusted approach allowed for inspection of individual subject variables associated with children who met criteria for SD and was parsimonious

relative to any claims about the theoretical or applied significance of the magnitude of the prevalence estimate. In the present context, if the prevalence obtained for speech delay had been adjusted for the percentage of false negatives obtained in the validation study, the estimated prevalence of SD in this sample would be 5.1% (i.e., increased by one third to adjust for the obtained sensitivity of 66.7; see Appendix). With a one-third increase in the comorbidity of SD and SLI indexed by either disorder, the approximately 5–15% comorbidity rates shown in Figure 3 and Figure 4 increase to approximately 6–20%. These rates are still notably lower than 19 of the 24 comorbidity estimates listed in Table 2, and lower than 7 of the 9 comorbidity estimates obtained for children of the same age or older.

A second possible explanation for the low comorbidity estimates compared to those reported in other studies is that they reflect increased efficacy of contemporary early childhood programs focusing on phonology (cf. Hodson, 1997). Many of the comorbidity estimates in Table 2 are based on samples of children who were actually assessed over a decade earlier than the children assessed in the present sample. The lower comorbidity estimates for more recent 6-year-old cohorts could, at least in part, reflect differences in the percentage of children enrolled in treatment programs and in the possible increased percentage of positive treatment outcomes. Notable as well are the number of girls with SD and low cognitive status (Figure 3), with cognitive needs increasing the likelihood of enrollment in early childhood programs. Some indication of the complex of factors associated with this explanatory perspective was presented elsewhere (Tomblin, 1996a, 1996b), indicating that the presence of a speech disorder significantly increases the probability of girls with language disorders being identified and enrolled for treatment. It is also notable in this regard that the prevalence obtained for SD (3.8%) by 6 years of age is nearly one half the prevalence of SLI (7.4%) reported in Tomblin et al. (1997) for the original sample. Thus, as previously discussed, normalization rates for speech versus language disorders—with or without early intervention—are likely to be significant moderators of prevalence and comorbidity data for children in this age range.

Implication for Genetics Studies

One of the goals of the present prevalence and comorbidity study was to relate findings to conceptual and methodological issues in emerging speech-genetics studies. Based on normalization rates, the present data suggest that liability estimates in speech-genetics research should be set considerably higher than 3.8% for preschool children and considerably lower than 3.8% for older children and adults.

In addition to use of the prevalence data for liability estimates and support for alternative transmission models, the present comorbidity findings have implications for phenotype issues. Essentially, the lower the comorbidity between SD and SLI, the less likely it is that these disorders may be transmitted by a common genotype and, therefore, that they should be defined with one common phenotype. Because the present data were obtained at 6 years of age, an age by which 75% of children with speech delay may have normalized, conjectures about genotype-phenotype relationships are expressly preliminary. Nevertheless, there are three methodological issues that invite brief consideration from the perspective of emerging speech-genetics studies.

Alternative Phenotypes for an Inherited Form of Speech Disorder

As discussed elsewhere, the concept of a speech *delay* reflects only one possible linguistic domain within the construct of atypical phonological development (Shriberg, 1993a, 1993b, 1994). Alternative phenotypes for an inherited form of child speech disorder can and have been based at many levels: psycholinguistic, descriptive linguistic, and manifest speech, with the last of these including alternative sampling contexts (e.g., citation speech, natural speech, challenging speech, intelligibility). Perhaps the most widely discussed candidate phenotype for a developmental phonological disorder is performance on one or more phonological processing tasks, such as segmentation, rhyming, or nonsense-word repetition tasks (cf. Edwards & Lahey, in press). Thus, as a candidate phenotype for speech-genetics studies, it is important to acknowledge the levels of speech processing for which the present definition and prevalence estimate for SD may or may not be sensitive because it is based solely on speech production. As disorder classifications or phenotypes are predicated on levels beyond manifest natural speech (e.g., deficits in phonological awareness), they invoke cognitive-linguistic levels that complicate the concept of a *speech* disorder. Accordingly, phonological disorder is becoming increasingly ambiguous as a cover term in both research and applied contexts.

Alternative Etiologies for Speech Delay

It is also important to note that genetic transmission of SD is only one of several hypothesized backgrounds for speech delay as defined and identified in this study. Other possible risk backgrounds of considerable interest to researchers include early recurrent otitis media with effusion, suspected developmental apraxia of speech, and issues associated with psychosocial development. Preliminary prevalence estimates

based on pedigree analysis of clinical referral samples suggest that familial SD may be suspected for as many as 60% of preschool children with speech delay (Shriberg, 1994; Lewis & Shriberg, 1994). Lahey and Edwards (1995) report a similar percentage (60%) for suspected familial SLI. Accordingly, all implications of the present findings for genetics research must be tempered by the possibility that different etiological backgrounds (i.e., *phenocopies*) may be associated with different values for variables such as prevalence estimates, sex ratios, normalization rates, and estimates of the comorbidity of speech delay with language impairment. For example, using the 60% estimate described above as a preliminary guideline for genetics studies, phenotypes extrapolated from the present prevalence and comorbidity data might apply to only 2.3% (i.e., $60\% \times 3.8\%$) of the children identified as SD in this study.

Alternative Epidemiologic Variables for Speech-Language Genetics Studies

Finally, the hypothesis of hereditary speech delay gains support to the degree that variables reflecting plausible environmental influences on speech acquisition are not associated or are only minimally associated with differential prevalence rates. In the present study, residential strata (urban, suburban, rural) is one such environmental variable that might appear to provide counterevidence for genetic influences because the prevalence of SD was generally higher in the urban stratum and lower in the rural stratum. However, residential strata is confounded with other sociodemographic variables that, in turn, affect the population genetics of children from similar racial/cultural backgrounds residing in the different strata. Genetic factors are likely moderated or mediated by a complex of environmental variables (i.e., multifactorial transmission models), with net effects again on prevalence, sex ratios, normalization outcomes, and comorbidity with SLI.

Conclusion

The present study was an attempt to use data available in the EPISLI project to estimate the prevalence of speech delay and its comorbidity with language impairment. The literature review and discussion have focused on theoretical and methodological issues relevant for the application of findings to emerging research attempting to identify the genotype underlying an inherited form of speech delay. The six primary findings, which will be subject to cross-validation in a forthcoming prospective study of a racially and culturally diverse preschool cohort, can be summarized as follows: (a) The prevalence of speech delay in 6-year-old children in this sample of 1,328 children was 3.8%; (b) at 6 years of age, speech

delay was approximately 1.5 times more prevalent in boys (4.5%) than girls (3.1%); (c) cross-tabulation by sex, residential strata, and racial/cultural background yielded prevalence rates for speech delay ranging from 0% to approximately 9% (with cell sizes posing a major reliability constraint for most estimates); (d) comorbidity of speech delay and language impairment was 1.3% (0.51% with SLI); (e) at 6 years of age, approximately 11–15% of children with persisting speech delay had specific language impairment (SLI); and (f) approximately 5–8% of children with persisting SLI had speech delay.

Acknowledgments

We extend our sincere thanks to the following people at the Iowa and Wisconsin sites for their expertise at the many stages of this study. In Iowa, the EPISLI project was coordinated by Paula Buckwalter, Marlea O'Brien, Nancy Records, and Xuyang Zhang. Speech-language and all other participant data were obtained by Chris Anderson, Kathleen Bailey, Jean Beisler, Lisa Ehlert, Connie Ferguson, Diane Highnam, Joni Mack, Chris McLaughlin, Jacqueline Nesvik, Julie Ann Sellen, Shirley Tiemeyer, Vickie Vandike, and Cathy Wignall. In Wisconsin, the research team included Chad Allen, Diane Austin, Sheryl Hall, Heather Karlsson, Carmen Rasmussen, Rachel Reily, Dorothy Ross, Carol Widder, and David Wilson. Peter Flipsen Jr. and Steve Pittelko provided useful comments on the manuscript. Copies of relevant technical reports are available at the Phonology Project Web site (www.waisman.wisc.edu/phonology). This study was supported by the following grants from the National Institute on Deafness and Other Communicative Disorders: NIH-DC-19-90, NIDCD DC02746, and NIDCD DC00496.

References

- American Psychiatric Association.** (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G.** (1986). Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *Journal of Speech and Hearing Disorders, 51*, 98–110.
- Bishop, D. V. M., & Edmundson, A.** (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*, 156–173.
- Caron, C., & Rutter, M.** (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychiatry, 32*(7), 1063–1080.
- Connell, P., Elbert, M., & Dinnsen, D.** (1991, June). *A syntax-delayed subgroup of phonologically delayed children*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
- Crowe Hall, B. J.** (1991). Attitudes of fourth and sixth graders toward peers with mild articulation disorders. *Language, Speech, and Hearing Services in Schools, 22*, 334–339.
- Edwards, M. L.** (1997). Historical overview of clinical phonology. In B. W. Hodson & M. L. Edwards (Eds.), *Perspectives in applied phonology* (pp. 1–18). Gaithersburg, MD: Aspen Publishers.
- Edwards, J., & Lahey, M.** (in press). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*.
- Elbert, M.** (1997). From articulation to phonology: The challenge of change. In B. W. Hodson & M. L. Edwards (Eds.), *Perspectives in applied phonology* (pp. 43–60). Gaithersburg, MD: Aspen Publishers.
- Enderby, P., & Philipp, R.** (1986). Speech and language handicap: Towards knowing the size of the problem. *British Journal of Disorders of Communication, 21*, 151–165.
- Fein, D. J.** (1983,). The prevalence of speech and language impairments. *Asha, 25*, 37.
- Felsenfeld, S., McGue, M., & Broen, P. A.** (1995). Family aggregation of phonological disorders: Results from a 28-year follow-up. *Journal of Speech and Hearing Research, 38*, 1091–1107.
- Fleiss, J. L.** (1981). *Statistical methods for rates and proportions*. New York: John Wiley & Sons.
- Flipsen, P., Jr., Shriberg, L. D., Weismer, G., Karlsson, H. B., & McSweeney, J. L.** (1999a). Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language, and Hearing Research, 42*, 663–677.
- Flipsen, P., Jr., Shriberg, L. D., Weismer, G., Karlsson, H. B., & McSweeney, J. L.** (1999b). *Acoustic characteristics of /r/ in adolescents*. Manuscript submitted for publication.
- Goffman, L.** (1998, June). *Prosodic influences on speech movements in children with Specific Language Impairment*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
- Himmelwright Gross, G., St. Louis, K. O., Ruscello, D. M., & Hull, F. M.** (1985). Language abilities of articulatory-disordered school children with multiple or residual errors. *Language, Speech, and Hearing Services in Schools, 16*, 171–186.
- Hodson, B. W.** (1997). Disordered phonologies: What have we learned about assessment and treatment? In B. W. Hodson & M. L. Edwards (Eds.), *Perspectives in applied phonology* (pp. 197–224). Gaithersburg, MD: Aspen Publishers.
- Hull, F. M., Mielke, P. W., Jr., Timmons, R. J., & Willeford, J. A.** (1971). The national speech and hearing survey: Preliminary results. *Asha, 13*, 501–509.
- Irwin, J. V., Huskey, R., Knight, N., & Oltman, S.** (1974). A longitudinal study of the spontaneous remission of articulatory defects of 1,665 school children in grades 1, 2 and 3. *Acta Symbolica, 5*, 1–17.
- Karlsson, H. A., & Shriberg, L. D.** (1999). *An acoustic description of /s/ productions in two subtypes of child speech disorders*. Manuscript in preparation.
- Kirkpatrick, E., & Ward, J.** (1984). Prevalence of articulation errors in New South Wales primary school pupils.

- Australian Journal of Human Communication Disorders*, 12(1), 55–62.
- Lahey, M., & Edwards, J.** (1995). Specific language impairment: Preliminary investigation of factors associated with family history and with patterns of language performance. *Journal of Speech and Hearing Research*, 38, 643–657.
- Lassman, F. M., Fisch, R. O., Vetter, D. K., & LaBenz, E. J.** (Eds.). (1980). *Early correlates of speech, language, and hearing*. Littleton, MA: PSG Publishing.
- Last, J. M.** (Ed.). (1988). *A dictionary of epidemiology* (2nd ed.). New York: Oxford University Press.
- Leske, M.** (1981). Prevalence estimates of communicative disorders in the United States: Speech disorders. *Asha*, 23, 217–228.
- Lewis, B. A., & Shriberg, L. D.** (1994, November). *Life span interrelationships among speech, prosody-voice, and nontraditional phonological measures*. Miniseminar presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA.
- McSweeney, J. L., & Shriberg, L. D.** (1995). *Segmental and suprasegmental transcription reliability* (Tech. Rep. No. 2). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.
- Morley, M. E.** (1965). *The development and disorders of speech in childhood*. Baltimore, MD: Williams and Wilkins.
- Newcomer, P., & Hammill, D.** (1988). *Test of Language Development–2 Primary*. Austin, TX: Pro-Ed.
- Newman, P. W.** (1961). Speech impaired? *Asha*, 3, 9–10.
- Paul, R.** (1993). Patterns of development in late talkers: Preschool years. *Journal of Childhood Communication Disorders*, 15, 7–14.
- Paul, R.** (Ed.). (1998). *Exploring the speech-language connection*. Baltimore: Brookes.
- Paul, R., & Shriberg, L. D.** (1982). Associations between phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research*, 25, 536–547.
- Peckham, C. S.** (1973). Speech defects in a national sample of children aged seven years. *British Journal of Disorders of Communication*, 8(1), 2–8.
- Prins, T. D.** (1962). Analysis of correlations among various articulatory deviations. *Journal of Speech and Hearing Research*, 5, 162–168.
- Roe, V., & Milisen, R.** (1942). The effect of maturation upon defective articulation in elementary grades. *Journal of Speech Disorders*, 7, 37–50.
- Ruscello, D. M., St. Louis, K. O., & Mason, N.** (1991). School-aged children with phonological disorders: Coexistence with other speech/language disorders. *Journal of Speech and Hearing Research*, 34, 236–242.
- Schery, T. K.** (1985). Correlates of language development in language-disordered children. *Journal of Speech and Hearing Disorders*, 50, 73–83.
- Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D.** (1990). Prevalence of reading disability in boys and girls: Results of the Connecticut Longitudinal Study. *Journal of the American Medical Association*, 264(8), 998–1002.
- Shriberg, L. D.** (1980). Developmental phonological disorders. In T. J. Hixon, L. D. Shriberg, & J. S. Saxman (Eds.), *Introduction to communicative disorders* (pp. 262–309). Englewood Cliffs, NJ: Prentice Hall.
- Shriberg, L. D.** (1982). Toward classification of developmental phonological disorders. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 8, pp. 2–18). New York: Academic Press.
- Shriberg, L. D.** (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Hillsdale, NJ: Lawrence Erlbaum.
- Shriberg, L. D.** (1993a). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*, 36, 105–140.
- Shriberg, L. D.** (1993b, May). *Phenotypes and phonologies: Facts and fictions*. Paper presented at the 14th Annual Symposium on Research in Child Language Disorders, Madison, WI.
- Shriberg, L. D.** (1994). Five subtypes of developmental phonological disorders. *Clinics in Communication Disorders*, 4(1), 38–53.
- Shriberg, L. D.** (1997). Developmental phonological disorder(s): One or many? In B. W. Hodson & M. L. Edwards (Eds.), *Perspectives in applied phonology* (pp. 105–127). Gaithersburg, MD: Aspen.
- Shriberg, L. D., & Austin, D.** (1998). Comorbidity of speech-language disorder: Implications for a phenotype marker for speech delay. In R. Paul (Ed.), *The speech-language connection* (pp. 73–117). Baltimore, MD: Paul H. Brookes.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L.** (1997a). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, 40, 708–722.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L.** (1997b). The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research*, 40, 723–740.
- Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J.** (1994). Developmental phonological disorders III: Long-term speech-sound normalization. *Journal of Speech and Hearing Research*, 37, 1151–1177.
- Shriberg, L. D., & Kwiatkowski, J.** (1980). *Natural process analysis: A procedure for phonological analysis of continuous speech samples*. New York: Macmillan.
- Shriberg, L. D., & Kwiatkowski, J.** (1982). Phonological disorders I: A diagnostic classification system. *Journal of Speech and Hearing Disorders*, 47, 226–241.
- Shriberg, L. D., & Kwiatkowski, J.** (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research*, 37, 1100–1126.
- Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B.** (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders*, 51, 140–161.
- Shriberg, L. D., Kwiatkowski, J., & Gruber, F. A.** (1994). Developmental phonological disorders II: Short-term

- speech-sound normalization. *Journal of Speech and Hearing Research*, 37, 1127–1150.
- Shriberg, L. D., & Lof, G. L.** (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics and Phonetics*, 5, 225–279.
- Shriberg, L. D., McSweeney, J., & Kwiatkowski, J.** (1999). *A prospective study of normalization in preschool children with speech-language delay*. Manuscript in preparation.
- Shriberg, L. D., & Olson, D.** (1988). *PEPAGREE: A Program to Compute Transcription Reliability*. Waisman Center Research Computing Facility, University of Wisconsin-Madison.
- Silva, P. A.** (1980). The prevalence, stability and significance of developmental language delay in preschool children. *Developmental Medicine and Child Neurology*, 22, 768–777.
- Silva, P. A., Justin, C., McGee, R., & Williams, S. M.** (1984). Some developmental and behavioral characteristics of seven-year-old children with delayed speech development. *British Journal of Disorders of Communication*, 19, 147–154.
- Silverman, F. H., & Paulus, P. G.** (1989). Peer reactions to teenagers who substitute /w/ for /r/. *Language, Speech, and Hearing Services in Schools*, 20, 219–221.
- Smit, A. B., & Bernthal, J. E.** (1983). Performance of articulation-disordered children on language and perception measures. *Journal of Speech and Hearing Research*, 26, 124–136.
- St. Louis, K. O., Ruscello, D. M., Grafton, S. J., & Hershman, K. T.** (1994, November). *Coexistence of communication disorders in clinical populations over four decades*. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA.
- St. Louis, K. O., Ruscello, D. M., & Lundeen, C.** (1992). *Coexistence of communication disorders in schoolchildren* (ASHA Monograph No. 27). Rockville, MD: American Speech-Language-Hearing Association.
- Stark, R. E., & Tallal, P.** (1988). *Language, speech, and reading disorder in children: Neuropsychological studies*. Boston: Little, Brown.
- Stevenson, J., & Richman, N.** (1976). The prevalence of language delay in a population of three-year-old children and its association with general retardation. *Developmental Medicine and Child Neurology*, 18, 431–441.
- Tallal, P., Ross, R., & Curtiss, S.** (1989). Familial aggregation in specific language impairment. *Journal of Speech and Hearing Disorders*, 54, 167–173.
- Tomblin, J. B.** (1996a). *The big picture of SLI: Epidemiology of SLI*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
- Tomblin, J. B.** (1996b, April). *Epidemiology of SLI: The association of speech sound disorder with SLI*. Paper presented at the Child Phonology Conference, Iowa City, IA.
- Tomblin, J. B., Records, N. L., Buckwalter, P. R., 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.
- Wallin, J. E. W.** (1916). A census of speech defectives among public-school pupils: A preliminary report. *School and Society*, 3(58), 213–216.
- Wechsler, D.** (1989). *WPPSI-R Manual: Wechsler Preschool and Primary Scale of Intelligence-Revised*. New York: Psychological Corporation.
- Whitehurst, G. J., Arnold, D. S., Smith, M., Fischel, J. E., Lonigan, C. J., & Valdez-Menchaca, M. C.** (1991). Family history in developmental expressive language delay. *Journal of Speech and Hearing Research*, 34, 1150–1157.
- Winitz, H.** (1969). *Articulatory acquisition and behavior*. New York: Appleton-Century-Crofts.
- Winitz, H., & Darley, F. L.** (1980). Speech production. In F. M. Lassman, R. O. Fisch, D. K. Vetter, & E. S. LaBenz (Eds.), *Early correlates of speech, language, and hearing* (pp. 232–265). Littleton, MA: PSG Publishing.

Received November 16, 1998

Accepted May 4, 1999

Contact author: Lawrence D. Shriberg, the Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, Wisconsin 53705. Email: shriberg@waisman.wisc.edu.

Appendix. Methods to Classify Speech Status Using TOLD-2: P Responses and the Speech Disorders Classification System (SDCS).

As developed in the text, speech delay (SD), as defined in the Speech Disorders Classification System (SDCS; Shriberg et al., 1997b), was proposed as the “gold standard” for the form of child speech disorders generally referred to as a developmental phonological disorder of unknown origin. To derive a prevalence estimate for SD from the EPISLI database, it was necessary to develop a procedure by which TOLD-2: P Word Articulation subtest responses could be validated to this standard. As discussed in the text, the age of children (5–6 years) in the EPISLI project was too young to estimate the prevalence of residual errors (RE), a classification in the SDCS for which eligibility begins at 9 years of age. However, the prevalence data would need to differentiate some children with questionable residual errors (QRE) on the SDCS from children with SD because eligibility for QRE in the SDCS approach begins at 6 years. The following sections summarize the several pilot studies, with information gained from each approach motivating the final procedure used to estimate the prevalence of SD in a sociodemographically balanced subset of 1,328 children from the EPISLI database.

Initial Pilot Study: Word Articulation Total Score

Tomblin (1996b) reported the first attempt to use speech classification data obtained from conversational speech samples in the EPISLI project to validate speech classifications based on TOLD-2: P Word Articulation (henceforth, WA) data. Tomblin's data set was a group of 164 children in the EPISLI project for whom WA data and conversational speech samples were available at that time. Methods for language and WA assessment are reviewed in Tomblin et al. (1997), and methods for speech sampling and analyses are reviewed in the current paper. The correlation between WA total scores and scores on the Percentage of Consonants Correct-Revised metric described in the present paper was .60, indicating approximately 36% common variance. A *z* score of -1.14 on the WA was determined to have optimal sensitivity and specificity to detect SD based on SDCS classification from conversational speech, although it yielded a false positive rate of .67.

Tomblin applied the -1.14 SD criterion to determine the comorbidity of SD in 1,927 children in the EPISLI database. Tomblin reported that SD cooccurred in 25% of children with SLI, but also in 18% of children with typically developing language. As anticipated by the false positive rate, the prevalence figure of 18% SD in children of this age with typically developing language was too high, suggesting an unacceptable specificity problem with the use of WA scores to discriminate SD. Clearly, high WA scores reliably detected children classified as normal speech acquisition (NSA) based on their conversational speech, but low WA scores yielded apparent false positives for SD.

Probable explanation of the unacceptable specificity was centered on the WA stimuli and method of scoring. Of the 20 WA test stimuli, 19 include one or more tokens of the phonemes /s/, /z/, /r/, /r̥/, or /æ̃/. Because distortions are scored as incorrect on the WA, children who only distort these

sounds will have very low WA scores. On the SDCS, children who only distort these sounds would be considered as having NSA or QRE, depending on whether they were younger or older than 6 years, respectively. Therefore, use of WA total scores to classify children's speech as SD versus NSA or QRE was eliminated as a sufficient method.

Second Pilot Study: Word Articulation Item Scores

The second pilot study was an item analysis of the WA responses of 82 children in the SLI project for whom both WA and conversational speech samples were available. The sample included 24 children with SD, 24 children with NSA, 10 children with an intermediate category between SD and NSA termed normal speech acquisition/speech delay (NSA/SD) (cf. Shriberg et al., 1997b), and 24 children whose distortion errors at this age qualified them for QRE status. (All speech classifications were based on SDCS output for the children's conversational speech samples.) The sample included data from 53 (65%) White children and 29 (35%) children of African-American, American-Indian, and Asian backgrounds. The first goal of this second pilot study was to determine if performance on a subset of items on the WA measure could be used alone or in combination with the total WA score to differentiate children whose speech in conversation was classified by the SDCS as SD, NSA/SD, QRE, or NSA. The second goal was to determine if the resulting procedures required any adjustments for speech features (i.e., dialectal variations) in children from different racial/cultural backgrounds.

A detailed item analysis confirmed the inherent problems in using these data to differentiate among the SDCS categories. There were no items or combinations of items that discriminated adequately among the primary categories. What did emerge was the need to disregard for each word certain errors that occurred for children classified as NSA or NSA/SD and were deemed acceptable for children of this age in the SDCS reference table (cf. Shriberg, 1993a, Appendix, Table A). For example, *basket* and *soldier* were challenging words for all children of this age and were frequently associated with consonant omissions and substitutions among children classified as NSA. Most distortion errors were considered acceptable at this age in the SDCS reference table, and, as suggested previously, most of the WA words included sounds on which distortions were frequent in children classified as NSA on the SDCS. There were three words on which substitutions and omissions of consonants consistent with African American English vernacular were also frequent. Although the item analysis did not yield a subset of words with high sensitivity and specificity for SD, it did yield a list of speech errors considered acceptable for each of the 20 WA test items based on the frequency of that error in children classified as NSA, QRE, or NSA/SD in the SDCS. Based on an array of developmental and transcriber-reliability criteria, alternative classification rules were also pretested for use in the criterion-validity study.

Criterion-Validity Study

The criterion-validity study included data from 521 children on whom conversational speech samples and resultant SDCS classification data were available from all associated research within the EPISLI project. Two research assistants coded the TOLD forms from these children using templates that reduced coding to a clerical task. The lead assistant rechecked all scoring completed by the other assistant, yielding no classification changes. A total of 58, or approximately 11%, of the 521 TOLD forms were uncodable due to missing information, most often for error types. This estimate of the codability rate (approximately 89%) was considered adequate for the prevalence study.

The following three classification rules referenced to Shriberg et al. (1997b, Table A2), together with the word-level adjustments described in the second pilot study were used to initially classify children as SD, QRE, or NSA. These rules reflect reliability constraints in transcription (e.g., less than two deletion/substitution errors are considered unreliable), as well as validity requirement for sensitivity and specificity needs in use of the TOLD data.

1. If there are more than two age-inappropriate deletion or substitution errors on the WA, classify as SD.
2. If there are two or fewer age-inappropriate errors, and more than six /r/, /r̄/, or /ə/ distortions and/or more than seven /s/ or /z/ distortions (i.e., distortions on more than 50% of opportunities on the WA), classify as QRE.
3. If there are two or fewer age-inappropriate errors, and fewer than six /r/, /r̄/, or /ə/ distortions and fewer than seven /s/ or /z/ distortions, classify as NSA.

Use of the word-level adjustments and the three rules yielded sensitivity-specificity rates ranging from 66.7% to 95.7% for the three SDCS classification categories. For the purposes of the prevalence estimate to follow, the only value deemed to be low was sensitivity for SD at 66.7%. Whereas

specificity problems apparent in the two pilot studies were adequately addressed by the WA word-adjustment and coding rules (SD specificity = 95.5%), these procedures were underestimating the occurrence of SD as validated to the SDCS conversational speech sampling procedures. Rather than adjusting for false negatives, as done in Beitchman et al. (1986), the approach selected for the prevalence study was to accept this level of sensitivity as a conservative perspective for the goals of the present study. Given the many conceptual issues associated with measurement of speech as a verbal trait, and especially the lack of an estimate of the standard error of measurement for the examiner's scoring of the WA, it seemed prudent to complete the prevalence study without adding unknown variance associated with a statistical adjustment. Also, as discussed in the text, due to the concerns about the validity and reliability of the QRE classifications from the TOLD data, prevalence and comorbidity data on this classification were deemed inappropriate to assemble and report. For the purposes of the present prevalence study no differentiation was made between NSA and QRE.

Prevalence Study

The sociodemographically balanced subsample of the EPISLI project described in the text originally included 1,456 children, including 303 children for whom conversational speech samples had been obtained and were classified using the SDCS program. Of the 1,456 children, 1,433 WA forms were available for processing at the Madison site. Of these 1,433 forms, 104 (7.3%) were ineligible for the coding procedure developed in the criterion-validity study. The codability rate of 92.7% of the available forms was considered adequate for the prevalence estimate. The prevalence estimates for SD and its comorbidity with cognitive-linguistic deficits were thus obtained on a subset of 1,328 children in the EPISLI project, including SDCS classifications for 303 children and WA results validated to SDCS classifications for the remaining 1,019 children.

Prevalence of Speech Delay in 6-Year-Old Children and Comorbidity With Language Impairment

Lawrence D. Shriberg, J. Bruce Tomblin, and Jane L. McSweeney
J Speech Lang Hear Res 1999;42;1461-1481

This article has been cited by 18 HighWire-hosted article(s) which you can access for free at:

<http://jslhr.asha.org/cgi/content/abstract/42/6/1461#otherarticles>

This information is current as of June 28, 2012

This article, along with updated information and services, is located on the World Wide Web at:

<http://jslhr.asha.org/cgi/content/abstract/42/6/1461>



AMERICAN
SPEECH-LANGUAGE-
HEARING
ASSOCIATION