Risk for Speech Disorder Associated With Early Recurrent Otitis Media With Effusion: Two Retrospective Studies

Lawrence D. Shriberg
Peter Flipsen Jr
Helen Thielke
Joan Kwiatkowski
University of Wisconsin–Madison

Marilyn K. Kertoy
University of Western Ontario
London, Ontario, Canada

Murray L. Katcher
Robert A. Nellis
University of Wisconsin–Madison

Michael G. Block
Starkey Laboratories, Inc.
Eden Prairie, MN

The goals of this two-part series on children with histories of early recurrent otitis media with effusion (OME) were to assess the risk for speech disorder with and without hearing loss and to develop a preliminary descriptive-explanatory model for the findings. Recently available speech analysis programs, lifespan reference data, and statistical techniques were implemented with three cohorts of children with OME and their controls originally assessed in the 1980s: 35 typically developing 3-year-old children followed since infancy in a university-affiliated pediatrics clinic, 50 typically developing children of Native American background followed since infancy in a tribal health clinic, and (in the second paper) 70 children followed prospectively from 2 months of age to 3 years of age and older. Dependent variables included information from a suite of 10 metrics of speech production (Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997a, 1997b). Constraints on available sociodemographic and hearing status information limit generalizations from the comparative findings for each database, particularly data from the two retrospective studies. The present paper reports findings from risk analysis of conversational speech data from the first two cohorts, each of which included retrospective study of children for whom data on hearing loss were not available. Early recurrent OME was not associated with increased risk for speech disorder in the pediatrics sample but was associated with approximately 4.6 (CI = 1.10–20.20) increased risk for subclinical or clinical speech disorder in the children of Native American background. Discussion underscores the appropriateness of multifactorial risk models for this subtype of child speech disorder.

KEY WORDS: diagnostic assessment, epidemiology, otitis media, phonology, speech disorders

Thirty years ago, Holm and Kunze (1969) reported that otitis media before age 2 was a risk factor for speech disorder at 5–9 years. Despite considerable research activity since publication of this widely cited case-control study, there currently is no consensus on the public health question: Are children with early recurrent middle-ear disease at increased risk for speech disorder? The goals of the present and companion (Shriberg, Friel-Patti, Flipsen, & Brown, 2000) paper are to address this question and to provide a preliminary descriptive-explanatory model relating early middle ear disease with or without hearing loss to later speech-language disorder. As described in the Method section of each paper, findings for both questions posed are viewed as

preliminary because they are based on data from three cohorts of children tested over a decade ago. The databases for these cohorts did not always provide for the types of sociodemographic and other information that are routinely collected in contemporary studies using epidemiologic methods. The unique contribution of the present reports are that they use contemporary speech production analysis metrics and a variety of statistical modeling techniques. For efficiency, methodological issues and procedures common to both papers are described only in this first paper. Also, full discussion of findings from the first paper is deferred to a summative perspective provided at the end of the second paper.

Review of Findings From Studies of Otitis Media With Effusion and Speech

Description

Table 1 is a summary of findings from 27 studies of otitis media with effusion and speech obtained from a literature search of publications in English during the last 30 years. Owing to definitional inconsistencies in the otitis media literature (cf. Daly, 1997), not all of the 27 studies define children as having effusion, nor do studies (including those to be reported in this paper) consistently differentiate among such terms as *frequent*, *recurrent*, *persistent*, or *chronic* otitis media (cf. Roberts, Burchinal, & Campbell, 1994). With some modifications in the definition of retrospective and ambispective designs, the format is consistent with information provided in a literature review completed over a decade ago by Roberts and colleagues (Roberts, Burchinal, Koch, Footo, & Henderson, 1988) and updated in Roberts, Burchinal, Davis, Collier, and Henderson (1991) and Roberts and Clarke-Klein (1994). Each study in Table 1 included at least one question about OME as a risk factor for speech disorder using one of three designs. Retrospective designs used otological records of OME history relative to current speech status. Ambispective designs used case records to establish otitis history, with participants then followed for a specified time period to monitor change. Prospective designs used data collection at one point in time (typically birth or soon thereafter) with repeated measures at controlled intervals to monitor otologic, impedance, and/or audiologic status, as well as speech status.

Of the 27 studies classified in the second column in Table 1, 13 (48%) were retrospective, 10 (37%) were ambispective, and 4 (15%) were prospective. A critical difference relative to the risk analysis methods used in the current studies was whether the participants in each study constituted a cohort, in which all or almost all members of a group of individuals were classified for exposure history (i.e., OME) and for outcome (i.e., speech disorder). Cohort designs are common in epidemiologic research where the goal is to calculate the risk of disease for all possible members of the cohort. Cohort studies are less prone to selection and measurement biases than case-control studies and are better able to establish the temporal relationship between exposure and disease. With reference to Table 1, 10 (37%) of the studies were cohort designs, with the remaining 17 (63%) using noncohort, case-control designs. Thus, of 27 otitis-speech studies conducted during the last 30 years, 48% have used retrospective designs (11% retrospective cohort plus 37% case-control) to test whether a positive history of OME is a significant risk factor for speech disorder.

The Participants columns in Table 1 indicate that otitis-speech studies have primarily been concerned with preschool-age children, with few outcome data available from children older than approximately 11 years. Cell sizes for OME and control groups have ranged from small groups of fewer than 10 children to large groups numbering several hundred children. Participant populations have reflected a variety of sampling methods, sociodemographic characteristics, and diagnostic classifications. OME status has been documented or estimated from many sources (primarily parental reports, medical records, and school reports) and assessment methods (primarily tympanometry, otoscopy, and history of insertion of pressure equalization tubes).

Findings

The right-most two columns in Table 1 indicate statistically nonsignificant and significant speech findings in these 27 studies. Two conclusions, based respectively on tallies for each column and methods in these studies, are that (a) there is limited evidence for a strong correlative association between early OME and concurrent or later speech disorder, and (b) there is no evidence for a direct causal association between OME and speech disorder. A total of 17 studies (63%) yielded negative findings (i.e., failure to reject a null hypothesis of no statistically significant speech difference associated with OME), whereas 21 studies (78%) yielded at least one statistically significant finding associating OME with deficits in one or more speech variables. Thus, 10 studies reported both negative and positive findings.

Methodological Critique

Critiques of the internal and external validity of the statistically significant findings in Table 1, as well as of findings in many other studies of OME and language, learning, and behavioral variables, primarily have focused on otologic, audiologic, and language-learning issues (e.g., Kavanagh, 1986; Lous, 1995; Paradise, 1997; Roberts, Wallace, & Henderson, 1997). General conclusions

Table 1 (page 1 of 2). Studies of otitis media with effusion reporting speech production outcomes.[a]

	Method		Participants			Ascertainment		Speech findings	
Authors	Design[b]	Type[c]	Age[d]	OME	Controls	Population	Documentation of OME	Nonsignificant[e]	Significant[e]
Holm & Kunze (1969)	R	Ca	5–9	16	16	Hospital outpatients	Parent report; medical records		T-D words
Needleman & Menyuk (Needleman, 1977)	R	Cr[f]/Ca	3–8	20	20	Details not reported	Medical records; school report		T-D words and sent. imitation
Lehmann et al. (1979)	R	Ca	2–6	42		Speech and language referrals (with OME)	Parent report; medical records		Mean DASE percentile = 22.2
Silva et al. (1982)	R	Co	5	47[g]	355	New Zealand child development study	Type A vs. B tympanogram; microscopic otoscopy		DASS
Shriberg & Smith (1983)	R	Ca	3–6	11 15	11 40	University clinic; delayed speech referrals	Parent report; PE tube insertions; tympanograms; audiograms[h]		Initial consonant and nasal changes
Schlieper et al. (1985)	A[i]	Ca	3–6	13	13	Pediatric referrals; OME history and mild conductive loss[j]	Medical records		Phonology errors[k] (spont. speech)
Hubbard et al. (1985)	R	Ca	5–11	24	24	Cleft palate patients	Early (0;3) vs. late (2;6) PE tubes		T-D words
Bishop & Edmundson (1986)	A	L/Ca	4 and 4;6	22	34	Language-disordered with and w/o OME	Parent report	PCC[l]; process use[m]; OME errors[n]	
Silva et al. (1986)	A	L/Co	5–9	39[o]	297	New Zealand child development study	Type A vs. B tympanogram; microscopic otoscopy		DASS (age 5) DAC (age 7 and 9)
Dyson et al. (1987)	R	Ca	3–5	20		Outpatient ENT clinic; daycare	Medical records	OME error II[p]	APP[q] OME error I[r]
Paden et al. (1987)	A	L/Co	1–3[s]	40[t]		Children scheduled for PE tubes	Medical records	No pre-tube factor predicts progress	Multiple factors predict progress
van der Vyver et al. (1988)	R	Ca	7–11	10	10	Spastic or ataxic C.P.; with and w/o OME	Details not reported		ATAA
Churchill et al. (1988)	R	Ca	3–6	15	15	Enrolled in therapy for speech delay	Parent report	8 other APP processes	Stridency deletion; cluster reduction[u]
Roberts et al. (1988)	P	L/Co	2–8	55		At risk for school failure[v]	Otoscopy; tympanometry	Cons. errors.[w] Total processes.[w] Ind. processes[w]	Time with OME x total process use after age $4\frac{1}{2}$
Paden et al. (1989)	A	L/Co	1–3[x]	14[y]	22[z]	Children with delayed speech; scheduled for PE tubes	Medical records	Initial use of cluster red.; Initial use of liq. dev.	3/5 processes at initial testing[aa] Initial and retest speech scores[bb]
Lous (1990)	R	Co	6–9	133[cc] 6[dd]	251 378	Danish schoolchildren; two towns	Parent report	SITO x OME history	SITO x current tymp. in better ear
Teele et al. (1990)	P	L/Co	7	141		Pediatric practice patients with and w/o OME	Otoscopy; tympanometry	Speech errors x OME duration before age 3	G-F (total score) x OME duration before age 3

Table 1 (page 2 of 2). Studies of otitis media with effusion reporting speech production outcomes[a].

	Method		Participants			Ascertainment		Speech findings	
Authors	Design[b]	Type[c]	Age[d]	OME	Controls	Population	Documentation of OME	Nonsignificant[e]	Significant[e]
Thielke & Shriberg (1990)	R	Co	3–6	14	14	Indian Head Start program enrollees	Medical records	PCC[ee]	Intelligibility Index[ff]
Lonigan et al. (1992)	A	L/Ca	4 and 5	20		Expressive language disordered	Medical records	T-D (words) at 4 x OM (any time period)	T-D (words) at 5 x OM at 18–24 months
Zargi & Boltezar (1992)[gg]	R	Ca	8–10	33[hh]	29	ENT patients with and w/o supp. OM	Parent report; medical records	Presence of artic. disorders	
Harsten et al. (1993)	P	L/Co	4 and 7	13	29	Swedish children; recruited at birth	Microscopic otoscopy; tympanograms; audiograms	Process use at age 4 and 7[ii]	
Manders & Tyberghein (1993)	A	L/Ca	4 and 5	18	40	Dutch children just prior to PE tubes	Details not reported	Presence of artic. disorders	
Paul et al. (1993)	A	L/Ca	3 and 4	8[jj] 12[kk]	13[jj] 11[kk]	Normal talkers vs. late-talkers	Parent report or known PE tube placement	G-F (at 3)[ll] TOLD (at 4)[mm]	
Hoey Hemmer & Bernstein Ratner (1994)	R	Ca	2–4	6	6	Same-sex dizygotic twins; discordant OME histories	Parent report; medical records	PAT percentiles G-F percentiles PCC[ee]; PFC[nn]	
Rvachew et al. (1996a)	A	L/Ca	9–18 months	9	9	Infants with early[oo] and late[pp] onset OME	Medical records	Vowel F1 freq.; F2 freq.; std. dev. of F1 freq.; F2/F1 ratio	Std. dev. of F2 freq. on vowels
Rvachew et al. (1996b)	A	L/Ca	6–12 months	9	9	Infants with early[oo] and late[pp] onset OME	Medical records		Canonical babble ratio; canonical utterance types
Abraham et al. (1996)	P	L/Co	2	8	8	Longitudinal infant study; with and w/o OME	Otoscopy	Front consonants in inventory; Final position consonant accuracy	Back consonants in inventory; Initial consonants in inventory; Initial position consonant accuracy; Process use

Note. ATAA = Afrikaans Test of Articulation Ability (South Africa); DAC = Dunedin Articulation Check (New Zealand); APP = Assessment of Phonological Processes; DASE = Denver Articulation Screening Exam; DASS = Dunedin Articulation Screening Scale (New Zealand); G-F = Goldman-Fristoe Test of Articulation; SITO = Staten Institute for Talelidende, Odense (Denmark); T-D = Templin-Darley; TOLD = Test of Language Development.

[a]Not all studies limited evaluation to "effusion" cases. [b]R = retrospective; P = prospective; A = ambispective. [c]Co = cohort; Ca = case control; Cr = cross-sectional; L = longitudinal. [d]Age at which speech production assessed (in years). [e]X indicates correlation between variables. [f]4 age groups (*n* = 5 each). [g]Participants with bilateral OME at time of testing. [h]At least one measure indicating OME. [i]Did 1 year follow-up but speech results not reported. [j]At initial testing (normal at 1 year follow-up). [k]Used checklist (no details given). [l]Percent Consonants Correct (single words). [m]As per McReynolds and Elbert (1981) (single words). [n]Changes in initial consonants and nasals (Shriberg and Smith, 1983). [o]Participants with bilateral OME at initial testing. [p]Initial consonant errors in 3/20 participants. [q]APP stimuli; 5 common processes + sound changes from Shriberg and Smith (1983); results delayed relative to normal. [r]Nasal errors in 9/20 participants (no significance test). [s]Tested every 3–4 months until no longer delayed or age 3. [t]Divided into 3 groups based on phonologic skill. [u]Includes /s/ clusters. [v]Socioeconomic criteria; 51/55 = African American. [w]No sign; correlation with OME overall (single words). [x]Tested every 3–4 months until age 4. [y]Still speech delayed at age 4. [z]Caught up to norms by age 4. [aa]Stridency deletion; velar deviations; postvocalic obstruent omissions (single words). [bb]Mean of 5 process scores. [cc]One or more OM episodes. [dd]Type B tympanogram (better ear) at time of testing. [ee]Percent Consonants Correct (conversational speech). [ff]Percent words understood by transcriber. [gg]Language spoken = Slovenian. [hh]Suppurative otitis only. [ii]Single word stimuli; overall ratings (0 = normal, 2 = deviant). [jj]Normal onset of first words. [kk]Late-talkers. [ll]Percentile ranks. [mm]Word articulation subtest score. [nn]Percent (place and manner) Features Correct in conversation. [oo]First treated for OME before age 6 months. [pp]No OME treatment before age 6 months.

are that no one study has assembled the appropriate methods in all relevant domains to test whether early, frequent otitis media with effusion is a risk factor for speech disorder. Detailed examination of the 27 speech-otitis studies in Table 1 suggests that differences among findings are likely associated with the complex of designs and methods used to address the question of OME as a risk factor for speech disorder.

The first source of support for this claim is the evidence shown in the lengthy set of footnotes for Table 1, expressly included to underscore differences in method across the 27 studies. These often crucial details have generally been lost to prior summative reviews. A second source of support for the claim of significant methodological diversity are findings from a comparative analysis of the 27 studies. Analysis of studies that yielded at least one significant finding versus those yielding at least one nonsignificant finding (thus some studies appeared in both groups) indicated that there was no one design or method variable that was clearly associated with significant or nonsignificant findings. Specifically, studies reporting at least one statistically significant finding were not more likely (a) to have used a particular research design (e.g., prospective, case-control), (b) to have included significantly more participants in the OME+ group (sample sizes averaged approximately 31 participants for both OME+ and OME– groups), (c) to have differed on the age at which outcomes were measured, or (d) to have used tympanometry versus otoscopy to validate the comparison group. Finally, of specific interest in the present context, studies reporting at least one significant finding were not more likely to have used a particular method for speech sampling (e.g., articulation tests, conversational speech samples) or speech analysis (e.g., percentage consonants correct, phonological process analysis).

Modeling the Short-Term and Long-Term Sequelae of OME

Applied perspectives on otitis-speech research concerns the time course of effects and effect size. The short-term and long-term consequences of OME for speech-language development might be posited to range from clinically insubstantial to clinically substantial involvement. In addition, moderating or mediating variables (cf. Baron & Kenny, 1986) include the age at onset; degree and duration of each episode; comorbidity with other disorders; and preventive, mitigating, or exacerbating environmental factors. Accordingly, speech production errors could be described as falling along a continuum of subclinical to clinical deficits for certain target sounds or subsets of sounds and on more global constructs such as intelligibility. Difficulties developing stable phonological representations might in turn be posited to have short-term or long-term consequences for rate and type of development in other linguistic and psychosocial domains. Specifically, what manifests as mild (i.e., subclinical) speech differences at one stage of phonological development might be causally associated with greater involvements in the same or other linguistic domains at later periods of development (i.e., transitive or "downstream" effects).

One explanatory route for downstream effects of early OME is *direct* association between the phonological deficit and the deficit in another domain. For example, unstable underlying representation of /h/ might lead to difficulties learning the pronoun system (e.g., *he, him, his, her*). Alternative *indirect* or *mediated* pathways invoke the cognitive-linguistic styles of information processing that may be engendered by the unstable speech signal, including inattention to the acoustic and visual cues in verbal learning (e.g., Feagans, Sanyal, Henderson, Collier, & Applebaum, 1987). Bishop and Edmundson (1986) provide an instructive perspective on effect size, mediation, and long-term sequelae:

> It is often assumed that the 14 to 40 dB hearing loss associated with otitis media (Bess, 1983) is not severe enough to affect language development. Yet this degree of hearing loss can impair speech discrimination (Dobie & Berlin, 1979; Mustain, 1979). Even if a child can discriminate speech sounds in a test situation, we cannot assume that the hearing loss is irrelevant for language learning. Rabbitt (1968) showed that if normal adults were asked to repeat a message, recall of the first part of the message was hampered if the second part was presented through white noise: the additional effort needed to listen to a noisy message seemed to use up so much concentration that they forgot the earlier part that had been heard clearly. The implication is that if one has to concentrate hard to perceive a degraded signal, then capacity for deeper language processing is reduced, so that a hearing loss that is not severe enough to affect ability to perceive language may nevertheless impair comprehension. (p. 322)

Roberts (1997), invoking concepts from dynamical systems theory, elaborates on Nozza's (1988) claim that even small disruptions of hearing may have amplified effects early in development:

> First, even mild hearing loss (average Sound Awareness Threshold = 16 dB) occurring periodically over time may have a measurable adverse effect on categorical responding by infants under specific input conditions....This position is consistent with the general principle of dynamic systems that small changes in relevant contextual

variables, particularly during transition periods, can result in dramatic changes in system behavior (Kelso, 1995; Thelen & Smith, 1994). (p. 514)

Summary and Overview of Research Design

There currently is no consensus on the question of whether early, frequent otitis media with effusion is a risk factor for speech disorder. Moreover, as reviewed in the companion paper, no study to date has provided a descriptive-explanatory account linking early OME with or without measures of hearing loss to later speech delay.

Rationale for the present report is based on the recent availability of a suite of speech metrics and lifespan reference data (Shriberg et al., 1997a, 1997b) that can be used to address the question of whether a history of significant otitis media with effusion places a child at increased risk for speech disorder. As reviewed in the following section, three databases assembled in the 1980s were available for this purpose. The two study samples in the present paper were a retrospectively sampled cohort of 35 typically developing children followed in a pediatrics clinic and a retrospectively sampled group of 50 children of Native American background followed in a tribal health clinic. Preliminary findings have been reported for all children in the pediatrics clinic sample (Shriberg, Kwiatkowski, Block, et al., 1984) and for 28 of the 50 children of Native American background (Shriberg, 1987; Thielke & Shriberg, 1990). Although these two data sets collected over a decade ago did not allow use of prospective assessment and lack information on children's hearing, they met methodological needs for studying early otitis media as a risk factor for speech delay. The primary purpose of the following paper in this series, which includes information on both otitis media and hearing, was to cross-validate risk findings in the present paper and to model associations among early otitis and hearing loss and later speech outcomes. A forthcoming report of a collaborative project will extend the findings in these reports, using data from a large epidemiologic study of children whose OME and hearing status were monitored from 2 months of age.

The research design for these studies is considered to successfully address four speech measurement constraints and one analysis constraint in prior OME speech studies. In the two retrospective studies to follow (a) speech samples were obtained from natural conversation, not from imitated or spontaneously evoked word or sentence tasks; (b) speech data were reduced using narrow-phonetic transcription, not by correct/incorrect scoring using broad-phonetic transcription; (c) speech acquisition was assessed with multiple metrics sensitive to different structural levels of phonology, not by one single index of speech competence; (d) statistical comparisons included information on subclinical-level involvement as well as clinical involvement; and (e) risk analysis techniques were used to provide clinically relevant quantitative estimates of the potential effects of early OME on productive speech.

Method

Study 1: Pediatrics Clinic Cohort

Participants

Classification. A computer search of records in a university-affiliated general pediatrics clinic in 1983 identified a cohort of 67 3-year-old children who had been followed at the clinic since birth. All children were from monolingual American-English-speaking homes and were free of major medical or psychosocial involvements. None of the children had received or been referred for speech-language services.

Under the direction of author MLK, a pediatrician at the clinic during this period, children's otologic histories were constructed from physician entries indicating episodes of otitis media, middle-ear effusion, retracted tympanic membrane, liquid behind the tympanic membrane, or resolving otitis media. The goal was to chronicle the duration of each episode of otitis media. Fourteen days were allowed for each reported episode of otitis media in one or both ears unless abnormal return visits occurred; in this event, the exact number of days of involvement was noted. This criterion was established by the pediatrician to coincide with antibiotic regimens and return visits. The procedure is conservative in relation to that of Teele, Klein, Rosner, and The Greater Boston Otitis Media Study Group (1984), who allowed 29 days for each episode of otitis media. Contact letters and follow-up phone calls to caregivers of the 67 children yielded a 52% sample of 35 children volunteered by their caregivers for a 2.5-hour test session to be conducted at the university hospital.

Description. The gender distribution in the volunteered sample, 18 (51%) boys and 17 (49%) girls, was proportional to the distribution of boys (52%) and girls (48%) in the total pediatrics clinic cohort. Mean age at testing was 3 years 8 months, with a standard deviation of 3.5 months. The 35 children came from parents with relatively high educational backgrounds; approximately 50% of the children had one or two parents who had completed a graduate degree. Although formal data on race/ethnicity were not obtained, examiner impressions of children and caregivers were that all participants assessed were white, from at least middle-class backgrounds, with General American dialects (including some regularized Wisconsin vowel changes).

Children's case records indicated from 0 to 58 total weeks of middle-ear involvement during the first 3 years of these children's lives. Most OME involvement occurred between 6 and 24 months of age. Plots of the mean weeks of middle-ear involvement from birth to 3 years also indicated comparable percentages of involvement for participating children compared to the total cohort. Thus, all comparisons supported a conclusion that the 35 volunteered children were representative of the total cohort of 67 children.

For the analyses to follow, the 35 children were assigned to one of two groups: OME+ or OME–. By age 3 years the 19 children in the OME+ group had more than 6 weeks of involvement and the 16 children in the OME– group had 0–6 weeks of involvement. The latter criterion for typical OME histories, equivalent to three or fewer episodes, is consistent with current methodology in OME research (e.g., Roberts, Gravel, Schwartz, Dollaghan, & Campbell, 1998). These two groups formed the *Total group* comparisons, which include all children in each study. Two other groups, termed *OME severity subgroup* comparisons, were constructed; they consisted of subsets of children selected by their severity of OME involvement. In Study 1, 9 children who had at least 18 weeks of involvement were included in the OME+ severity subgroup. The range of involvement was 18 to 58 weeks (M = 26.4, SD = 12.8). A comparison OME– group was assembled of 9 children with 0 to 2 weeks of involvement (M = 1.1, SD = 1.1).

Table 2 includes descriptive data for Study 1 children, as well as for children in the other study group (Study 2) to be reported. As shown in the data columns for Study 1, Total group and OME severity subgroup, the two middle-ear involvement subgroups did not differ significantly in composition by gender or age. Three years of age might be considered too early to assess potential long-term effects of OME on speech, language, and learning domains. However, this age is maximally sensitive to the effects of OME on speech production before uncontrolled sources of normalization in older children obscure relevant error patterns. None of the Study 1 children in either group had received or been referred for speech-language services.

Otologic, audiologic, and language characteristics. All testing was conducted by two experienced speech-language examiners and the pediatrician, each of whom was unaware of a child's OME history. The two speech-language pathologists were randomly assigned to test approximately half of the children; the pediatrician examined all of the children. Each speech-language pathologist completed all tests in one of two sound-proof booths. The protocol included a standard examination of the oral mechanism, pure tone threshold testing at 500, 1000, and 2000 Hz (ANSI, 1969), acoustic-immittance measures (Grayson Stadler Auto Tymp 28), and clarification and discussion of items on a standard case history form completed by the caregiver.

On otoscopic examination, 87% of the OME– children had at least one otoscopically normal ear, compared to 67% of the OME+ children, but the difference was not statistically significant [$\chi^2(1, N = 33) = 1.782, p > .05$]. The two groups did not differ by pure tone averages for

Table 2. Gender and age characteristics of children with (OME+) and without (OME–) histories of OME in Study 1 and Study 2, including information on OME severity subgroups assembled for additional analyses.

	Total group			OME severity subgroup		
	OME+	OME–	*p*	OME+	OME–	*p*
Study 1						
n						
Boys	10	8	<.88[a]	2	5	<.15[a]
Girls	9	8		7	4	
Total	19	16		9	9	
Age (months)						
M	43.7	43.6	<.93[b]	43.6	42.9	<.73[b]
SD	3.6	3.7		3.9	4.3	
Study 2						
n						
Boys	9	16	<.77[a]	8	8	<1.00[a]
Girls	10	15		6	6	
Total	19	31		14	14	
Age (months)						
M	55.8	57.5	<.43[b]	56.5	57.6	<.69[b]
SD	7.2	7.2		7.2	7.0	

[a]χ^2 test. [b]Two-sample *t* test.

either the left ear [$t(29) = 0.29$, $p > .20$] or right ear [$t(32) = 0.51$, $p > .20$]. Acoustic-immittance data were suggestive of more marginal tympanograms (flat tympanograms; peak pressures less than –50 or greater than +50 daPa) for the OME+ group, but Fisher exact tests of proportions for each ear were nonsignificant (right ear: $p = .49$; left ear: $p = .09$). There were no significant between-group differences or trends on the speech perception and language protocol, which included the following measures commonly used in the early 1980s: Northwestern University Children's Perception of Speech (Elliott & Katz, 1980), Peabody Picture Vocabulary Test–Revised, Form M (Dunn & Dunn, 1981), Miller-Yoder Test of Linguistic Comprehension (Miller & Yoder, 1984), Illinois Test of Psycholinguistic Abilities Grammatic Closure Subtest (Kirk, McCarthy, & Kirk, 1968), an elicited-sentence procedure used to assess syntax, and a 5-min conversational sample used to compute mean length of utterance. Additional details on the language measures are available in Kertoy and Shriberg (1984).

Study 2: Children of Native American Background

Participants in the second study were 50 monolingual American English-speaking children of Native American background living on the Menominee Indian Reservation in Wisconsin. The prevalence of OME is high among Native Americans (Raymond, Garcia, & Scheib, 1993; Stewart, 1986; Toubbeh, 1985), particularly in this tribe (Goinz, 1984). A total of 28 of the children were those described in Thielke and Shriberg (1990). A second group of 22 children were added to this group to increase the sample size. The following sections describe the sampling and assessment procedures for the two subsamples.

First Subsample ($n = 28$)

Classification. A two-phase procedure was used to classify the otitis media histories of each of the 140 3- to 6-year-old children who were enrolled in the Head Start program on the Menominee Reservation in 1986. All 140 children were monolingual American English speaking and from similar socioeconomic backgrounds reflecting the 90% disadvantaged criterion required for Head Start placement on the reservation. Examiner impressions indicated that all eventual participants spoke General American dialect (including some regularized Wisconsin vowel changes).

First, Head Start student health and educational records were reviewed in consultation with the Health Coordinator and the Education/Handicapped Coordinator. Children who met one or more of the following criteria were eliminated from the pool of participants: suspected or diagnosed as having developmental disabilities, had not received recommended correction for vision problems, and had not received medical care at the tribal health clinic. This phase eliminated 22 potential participants, reducing the candidate pool to 118.

Second, records of hearing screenings from three sources were inspected: routine school screenings, records from a special Childhood Audiology Project operating in a 2-year period before the intended study, and pediatric procedures in the tribal health clinic. The special project had screened children's hearing on a quarterly basis (see Thielke, 1988, pp. 35–37, for complete description of otoscopic, audiologic, and impedance measures used in this project). The medical records at the tribal health clinic were reviewed in detail to tally the occurrence and types of middle-ear disease observed in well-baby examinations, medical treatments for specific illnesses, audiological evaluations, and referrals for suspected otitis media. A diagnosis of *acute otitis media* was based primarily on pneumatic otoscopy, and a diagnosis of *otitis media with effusion* was based on tympanometry. Records indicated that 80% of the diagnoses and treatments were provided by the same physician.

The goal of the two-phase classification procedure was to select approximately 30 children for testing who would differ maximally by OME history and be best matched in gender, age, place of residence on the reservation, and history and type of educational intervention services received before and during attendance at the Head Start program. The last variable included participation in an infant-stimulation program for children under age 3 years. Additional details on all phases of the records search are provided in Thielke (1988, pp. 37–40). The medical records review yielded a rank-ordering of the 118 children in terms of severity and chronicity of OME episodes, including chronological data establishing absence or presence of otitis media. From this list a total of 28 participants were eventually assigned to two groups of 14 participants each. The number of medical treatments for middle-ear disease for the OME+ group ranged from 6 to 23, whereas histories for children in the OME– group included 0–1 medical treatments. The average months of educational intervention services received by the 14 OME+ children ($M = 4.4$ months, $SD = 9.6$) and the 14 OME– children ($M = 2.6$ months, $SD = 3.2$) in the first Study 2 subsample were not significantly different [$F(1, 26) = .15$, $p > .05$]. OME severity subgroups were also assembled for these 28 children. Their medical records allowed this level of detail, which was not possible for the 22 children (to be described) added to constitute the Total group. As shown in the Total and OME severity subgroup comparisons in Table 2, OME+ and OME– subgroups assembled for the speech analyses did not differ significantly in gender or age.

Assessment. Two experienced examiners administered an assessment protocol to the 28 children in three sessions within a 2-week period. In the first session, an audiologist obtained audiologic and acoustic-immittance data. In the second and third sessions, a speech-language pathologist completed the remainder of the protocol, including the Photo Articulation Test (Pendergast, Dickey, Selmar, & Soder, 1969), a 10- to 15-minute conversational speech sample following the procedures described in Shriberg and Kwiatkowski (1980), the Peabody Picture Vocabulary Test–Revised, Form M (Dunn & Dunn, 1981), and subtests 5 and 6 from the Test of Auditory Comprehension (Trammel, 1977). Scores on the performance subtest of the Wechsler Pre-School and Primary Scale of Intelligence (Wechsler, 1967), which had been administered for other purposes by a school psychologist, were available for 18 of the 28 participants. The conversational speech samples were obtained and transcribed by two research transcribers using the same model recorders and recording techniques as used in Study 1.

Otologic, audiologic, and language characteristics. Based on pass/fail criteria of 0.2 to 1.8 ml compliance values and –200 to 100 mm H_2O pressure values, proportionally more children in the OME+ group failed the peak compliance [$\chi^2(1, N = 28) = 4.76, p < .05$] and peak pressure [$\chi^2(1, N = 28) = 4.76, p < .05$] tests computed for both ears for each participant. Participants in both groups failed the pure tone screening at 15 dB HL; however, all passed at 25 dB HL, with nonsignificant differences between the groups [$\chi^2(1, N = 28) = 0.7, p > .05$] computed on a pass/fail criterion for both ears for each participant. There were no statistically significant differences between the two middle-ear status groups in nonverbal intelligence [$F(1, 26) = 0.29, p > .05$]. Data for the Peabody Picture Vocabulary Test, analyzed by three scoring alternatives, yielded nonsignificant differences in age equivalent scores [$F(1, 26) = 1.23, p > .05$], but significantly lower scores for the OME+ group were found when scores were transformed to percentile [$F(1, 26) = 11.26, p < .05$] and stanine [$F(1, 26) = 8.73, p < .05$] equivalents. Comparison of the two subtests of the Test of Auditory Comprehension, requiring auditory memory for two critical elements (TAC 5) and four critical elements (TAC 6) in sentential material, also yielded significantly lower scores for children in the OME+ group on both measures [TAC 5: $F(1, 26) = 9.06, p < .05$; TAC 6: $F(1, 26) = 8.26, p < .05$].

Second Subsample (n = 22)

In 1990 and 1991, conversational speech samples from 22 children assessed by the speech-language pathologist who tested the first group of children were added to this database. Before speech sampling, the examiner (author HT) for the original study spent 2 days visiting and interacting in play activities with a group of approximately 145 children attending the Head Start program at two sites on the reservation. As an associated goal of these visits was to obtain normative data on a group of speech-language measures, all children who had received speech-language services or had been referred for any exceptional education need were eliminated from further consideration. The remaining children were then randomly asked to participate in the normative study; data for the present study were obtained from the children who volunteered to come to the examiner's room to talk about pictures.

When samples from a total of 22 randomly selected children were collected for the purpose of the present study, using the same speech assessment protocol as used in the prior study, the examiner conducted a thorough review of subject records in the tribal health clinic. The record review included information on the 22 children's health and developmental history, vision and hearing screenings, physical examinations, and teacher progress reports. These data indicated that five children had OME histories consistent with histories of children classified as OME+ in the first subsample, with the remaining 17 children classified as OME–. All speech sampling and transcription instrumentation and procedures were similar to those reported above. As shown in Table 2, there were no statistically significant differences in the gender or age composition of the 19 OME+ and 31 OME– children constituting the Study 2 Total group.

Speech Measurement

Procedures

A conversational speech sample was obtained for each child in each of the two studies following procedures described in Shriberg and Kwiatkowski (1980). The 10- to 15-min conversational samples, including at least 70 utterances and 90 word types, were used for both speech and language analyses. Sony 5000 audiocassette recorders, matching external microphones, and high-quality audiocassette tapes were used, with microphone-to-mouth distance monitored at approximately 15 cm. Although Study 1 recording was accomplished in a sound-proof booth and Study 2 samples in a quiet room, both recording contexts yielded tape recordings with excellent signal-to-noise characteristics.

Using Dictaphone Model 2550 transcription devices, two of the authors (Study 1) and two other experienced transcribers (Study 2) transcribed all samples following a system of narrow-phonetic transcription and conventions developed for research in child phonology (Shriberg & Kent, 1982; Shriberg, Kwiatkowski, &

Hoffmann, 1984). The transcribers were blind to all children's OME status. Samples were formatted for computer analysis using enhancements to the PEPPER system (Shriberg, 1986, 1993). The program provided scores on each of the 10 measures of articulatory competence described in Shriberg et al. (1997a). Nine of the 10 speech metrics treat articulatory competence as a criterion-referenced continuous trait, with a score of 100% on a metric reflecting maximum competence. The 10th measure provides a categorical classification of a child's speech status using a hierarchical polychotomous typology (cf. Shriberg, 1993; Shriberg & Austin, 1998; Shriberg et al., 1997b).

Measures

The 10 speech measures are defined and calculated as follows.

1. The *Percentage of Consonants Correct* (PCC) is a measure of the percentage of intended consonants produced correctly, with all deletions, substitutions, and clinical distortions (cf. Shriberg, 1993, Appendix) counted as incorrect. Subscales for consonants divided into three developmental sound classes—termed the Early-8, Middle-8, and Late-8 consonants—are available for the PCC (as well as for the PCC-A, PCC-R, and PCI described below). Percentage calculations for the three subscales are completed in the same way as described for the Total score on the PCC. The developmental sound class subscales provide measurement sensitivity to consonant sounds that children with significant speech disorder usually (Early-8), sometimes (Middle-8), and seldom (Late-8) articulate correctly.

2. The *Percentage of Consonants Correct–Adjusted* (PCC-A) is calculated in the same way as the PCC, except that five speech-sound distortions termed the *common clinical distortions* (cf. Shriberg, 1993, Appendix) are scored as correct.

3. The *Percentage of Consonants Correct–Revised* (PCC-R) is also calculated in the same way as the PCC, except that *all* clinical distortions are counted as correct.

Notice that these two alternatives to the PCC—the PCC-A and PCC-R—are sensitive to differences in age and speech characteristics of different speaker groups. Briefly, the PCC-A nullifies the attenuation of true differences in speech competence among speakers by removing the contribution of common clinical distortions to the numerator in the competence calculations. Similarly, the PCC-R nullifies the attenuation effects of all speech-sound distortions on the measurement of competence, reflecting only the percentage of speech-sound deletions and substitutions.

4. The *Percentage of Consonants in Inventory* (PCI) is a measure of the percentage of the 24 consonants that are articulated correctly at least once in conversational speech (cf. Shriberg et al., 1997a, for computational procedure).

5. The *Percentage of Vowels/Diphthongs Correct* (PVC) is calculated in the same way as the PCC, reflecting the percentage correct of all intended American English vowels and diphthongs.

6. The *Percentage of Vowels/Diphthongs Correct–Revised* (PVC-R) is calculated in the same way as the PCC-R, with all distortions of vowels and diphthongs scored as correct.

7. The *Percentage of Phonemes Correct* (PPC) is calculated in the same way as the PCC and PVC, combining the two values to yield one score reflecting articulatory competence on all English phonemes.

8. The *Percentage of Phonemes Correct–Revised* (PPC-R) is calculated in the same way as the PCC-R and the PVC-R, providing a score that reflects only deletion and substitution errors on all consonants and vowels/diphthongs in the speech sample.

9. The *Intelligibility Index* is derived from the percentage of child-intended words in the conversational speech sample that the examiner and/or transcriber could gloss. The Intelligibility Index reflects a best-case estimate of a speaker's intelligibility, because the sampling procedures require the examiner to provide a verbal gloss of each utterance on the tape and because transcribers are encouraged to use multiple playbacks to attempt a gloss of difficult strings.

10. The 10th speech measure computed for all children was status on the *Speech Disorders Classification System* (SDCS; Shriberg, 1993; Shriberg et al., 1997b). The SDCS program sorts a transcript of a conversational speech sample into one of several age-relevant clinical classifications, including for the present speakers (a) Normal (or Normalized) Speech Acquisition (NSA); (b) Normal (or Normalized) Speech Acquisition/Speech Delay (NSA/SD), an intermediate classification between normal and delayed speech; and (c) Speech Delay (SD). The SDCS program assigns classifications by tallying all speech errors in a transcript and, using a series of validity and reliability conventions (Shriberg et al., 1997b), comparing those errors to a table of developmental data assembled from the child phonology literature (cf. Shriberg, 1993, Appendix).

Transcriber Agreement

Point-by-point percentage of agreement data for the transcribers involved in these studies have been reported in detail, including information on broad- and narrow-phonetic transcription of consonants and vowels/diphthongs (McSweeny & Shriberg, 1995). For a representative sample of 32 conversational speech samples,

interjudge percentage of agreement for consonant transcription ranged from 90.0% to 95.1% (broad transcription) and from 73.9% to 85.3% (narrow transcription); for vowels/diphthongs, the range was 85.7–97.3% (broad transcription) and 71.2–85.3% (narrow transcription). Standard error of measurement estimates for the first nine speech measures above have also been reported, averaging 1–3 percentage points for the total scores and 3–6 percentage points for the consonant subscales (Shriberg et al., 1997a, Table 3). Prior estimates of the reliability of intelligibility index findings indicate intrajudge and interjudge agreement ranging from approximately 70% to 100%, depending greatly on a host of subject and sampling issues (cf. Kwiatkowski & Shriberg, 1992; Shriberg et al., 1997b; Shriberg & Kwiatkowski, 1980, 1982; Weston & Shriberg, 1992).

Results

Statistical Approach

Psychometric review of the speech measures data (e.g., cell sizes, skew and kurtosis, standard deviation ratios, correlations between means and standard deviations, percentage of 100% scores) suggested that these data did not always meet assumptions for parametric analyses. Therefore, means and standard deviations were used for descriptive statistics, but nonparametric Wilcoxon-Mann-Whitney rank order statistics (Siegel & Castellan, 1988) were completed for all between-group comparisons when measures were treated as continuous variables.

The large number of statistical tests within and across questions requires a rationale for setting the alpha levels required for statistical significance. Given the theoretical and applied goals, it was considered equally important to avoid both Type I and Type II errors. Rather than using arbitrary family-wise criteria to set significance levels (i.e., Bonferroni corrections), the decision was to acknowledge all obtained *p* values at the .05 level or less as statistically significant and to seek replicated findings across analyses and dependent variables. Lahey and Edwards (1995) provide a well-reasoned rationale for using an even more liberal alpha level of .10 in a study with similar measurement constraints and descriptive-explanatory goals. Lahey and Edwards argue as follows:

> This study is exploratory in nature and is meant to generate rather than to test hypotheses. Because of this, we have reported all probability levels, allowing the reader to judge whether differences noted are worthy of further exploration....By setting the alpha level at such a high value, we decrease the likelihood of missing differences that may be worthy of future study. (p. 644)

Between-Group Analyses

Table 3 includes descriptive and inferential statistics comparing children in each of the OME status groups in Study 1 and Study 2 on each of the nine quantitative speech indices. For each study, data are summarized for two analyses, the *Total group* comparisons and the *OME severity subgroup* comparisons. Power estimates for all Wilcoxon-Mann-Whitney median comparisons were obtained with the software program PASS (Hintze, 1996), which uses an approximation of the 2-sample *t* test in which the sample size is adjusted as suggested by Al-Sunduqchi (1990). A normal distribution was assumed, which resulted in a sample size adjustment factor of $n/(\pi/3)$. Power values reflect 1-Beta, with Beta being the risk of a Type II error. Power values for Study 1 ranged from .58 to .05 indicating that, given the effect sizes observed and the sample sizes used, the risk of a Type II error ranged from 42% to 95%. For study 2, with larger sample sizes, power values ranged from .87 to .09, indicating that the risk of a Type II error ranged from 13% to 91%.

Beginning with the Study 1 results in Table 3, none of the Total group comparisons for the nine speech metrics and their subscales (total = 20 comparisons) was statistically significant at the .05 alpha level. For the Study 1 OME severity subgroup comparisons, 3 of the 20 comparisons (15%) were statistically significant. Children with the most severe OME+ histories scored significantly higher than children with essentially no histories of OME (OME–) on the PCC Early-8 ($p = .025$), PCC Middle-8 ($p = .028$), and PCC-A Middle-8 ($p = .029$) consonants. Other than chance occurrence, there is no explanation for these counterdirectional findings.

For Study 2, 12 of the 20 Total group statistical comparisons (60% of the tests) indicated that the speech of children with OME+ histories was significantly less articulate than the speech of children with OME– histories. Similar directional findings were obtained for the OME severity subgroup comparisons, for which 9 of the 20 comparisons (45%) indicated significantly less articulate speech in the subgroup of children with the most severe OME+ histories. As each of the measures reflects a different perspective on speech competence, the findings in Table 3, Study 2, for Total group children are summarized in the following six paragraphs.

1. *Consonants correct in conversational speech.* For the Study 2 Total group analyses, as shown in Table 3, PCC scores for the OME+ group averaged exactly 5 percentage points lower than scores for the OME– group (80.6% and 85.6%, respectively) ($p = .024$). Thus, children in Study 2 with OME+ histories had fewer consonants correct in conversational speech when speech errors are defined as deletions, substitutions, and distortions of consonants. Between-group differences were in the same

Table 3. Speech comparisons for children with (OME+) and without (OME–) histories of OME in two retrospective cohort studies (Study 1, Study 2). In each study, between-group comparisons are presented for total groups and for a subgroup of children with the most and least severe OME histories.

		Total group					OME severity subgroup				
		OME+		OME–			OME+		OME–		
		M	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*
Study 1											
PCC	Early	98.1	2.1	97.3	3.3	0.436	99.1	0.9	97.0	2.9	0.025*
	Middle	92.4	11.5	95.0	3.3	0.583	97.2	1.7	94.7	2.5	0.028*
	Late	47.5	15.9	48.5	17.8	0.766	50.9	19.0	46.5	19.6	0.596
	Total	80.9	6.5	82.0	5.9	0.518	83.2	6.6	81.3	6.9	0.691
PCC-A[a]	Middle	92.5	11.5	95.1	3.2	0.560	97.2	1.7	94.8	2.4	0.029*
	Late	78.3	13.5	79.9	9.7	1.000	82.1	10.4	79.1	11.5	0.566
	Total	90.6	6.1	91.4	4.5	0.727	93.3	3.2	91.1	4.5	0.352
PCC-R	Early	98.7	1.7	98.3	1.6	0.328	99.4	0.5	98.1	1.8	0.051
	Middle	93.0	11.4	95.9	3.3	0.920	97.9	1.5	95.9	2.7	0.087
	Late	80.4	12.9	81.2	9.8	0.961	84.1	9.2	80.5	11.3	0.426
	Total	91.6	5.8	92.5	3.7	0.829	94.3	2.7	92.3	3.9	0.287
PCI	Early	100.0	0.0	100.0	0.0	—	100.0	0.0	100.0	0.0	—
	Middle	90.3	13.2	95.6	5.8	0.285	97.2	4.2	94.0	6.6	0.380
	Late	90.9	12.7	89.7	12.8	0.831	94.5	11.0	90.2	13.3	0.688
	Total	94.3	5.9	95.3	5.1	0.557	97.5	3.0	95.0	5.2	0.469
PVC		93.1	2.4	92.5	5.4	0.934	93.7	1.6	92.0	6.8	1.000
PVC-R		98.0	1.8	97.2	3.6	0.697	98.8	1.5	96.3	4.5	0.113
PPC		85.8	4.4	86.2	5.0	0.666	87.4	4.1	85.6	5.9	0.658
PPC-R		94.2	3.9	94.4	3.2	0.881	96.1	1.7	93.9	3.6	0.181
Int. Index		89.1	8.3	93.8	3.7	0.107	90.9	6.5	94.2	3.6	0.177
Study 2											
PCC	Early	95.9	3.1	97.1	2.8	0.151	95.3	3.5	96.3	3.3	0.564
	Middle	88.2	8.8	92.0	4.9	0.152	86.2	9.4	90.7	5.1	0.279
	Late	51.2	17.5	63.1	16.0	0.023*	43.7	12.7	52.7	15.6	0.103
	Total	80.6	7.8	85.6	6.0	0.024*	77.4	6.2	81.3	5.4	0.103
PCC-A	Middle	88.2	8.8	92.1	4.9	0.151	86.3	9.5	90.9	5.1	0.267
	Late	72.3	10.8	79.0	8.5	0.021*	69.3	9.4	78.4	9.2	0.010*
	Total	86.8	5.5	90.5	3.5	0.017*	85.0	5.0	89.3	3.4	0.032*
PCC-R	Early	96.6	2.9	97.9	2.6	0.245	96.0	3.2	96.7	3.3	0.907
	Middle	89.7	7.5	92.8	4.6	0.160	88.0	7.9	91.5	5.0	0.258
	Late	73.6	10.1	80.4	8.6	0.014*	70.8	8.8	80.4	9.0	0.004**
	Total	87.9	5.0	91.4	3.4	0.011*	86.3	4.4	90.3	3.4	0.019*
PCI	Early	99.7	1.4	100.0	0.0	—	99.6	1.7	100.0	0.0	—
	Middle	94.7	6.0	97.7	3.8	0.053	94.2	6.7	98.5	3.1	0.047*
	Late	87.0	10.4	92.5	7.3	0.071	86.6	11.2	90.8	7.6	0.351
	Total	94.1	3.3	96.9	2.6	0.005**	93.8	3.6	96.5	2.6	0.045*
PVC		94.5	3.2	96.7	1.9	0.018*	93.6	3.2	95.3	1.5	0.156
PVC-R		97.6	2.0	99.1	0.7	0.003**	97.5	2.3	99.1	0.7	0.034*
PPC		86.2	5.8	90.1	4.2	0.023*	83.9	4.8	86.9	3.5	0.129
PPC-R		91.8	3.6	94.5	2.2	0.008**	90.7	3.4	93.8	2.1	0.021*
Int. Index		96.5	3.1	97.8	2.4	0.041*	96.2	3.3	98.9	1.0	0.002**

Note. Table 2 provides *n*s for the OME+ and OME- subgroups for Study 1 and Study 2. *p* values are from the Wilcoxon-Mann-Whitney tests. Dashes indicate that no statistical test was possible. PCC = Percentage of Consonants Correct; PCC-A = Percentage of Consonants Correct–Adjusted; PCC-R = Percentage of Consonants Correct–Revised; PCI = Percentage of Consonants in the Inventory; PVC = Percentage of Vowels Correct; PVC-R = Percentage of Vowels Correct–Revised; PPC = Percentage of Phonemes Correct; PPC-R = Percentage of Phonemes Correct–Revised; Int. Index = Intelligibility Index. [a] values for PCC-A Early-8 consonants are the same as those for PCC Early-8 consonants.

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

direction for the OME severity subgroup comparisons, but the approximately 4-percentage point difference for these fewer participants was not statistically significant (p = .103).

2. *Consonant deletion and substitution errors.* As shown in Table 3, Study 2, children with OME+ histories had significantly lower scores on the PCC-A Late-8, PCC-A Total, PCC-R Late-8, and PCC-R Total in both the Total group comparisons and in the OME severity subgroup comparisons. These findings indicate that children in Study 2 with OME+ histories had significantly more consonant deletion and substitution errors than the children with OME– histories. Recall that in comparison with the PCC, the PCC-A scores common clinical distortions as correct and the PCC-R scores all clinical distortions as correct.

3. *Consonant inventories.* The fourth metric in Table 3, the PCI, is an index of the percentage of the 24 English consonants that were attested as *in* a child's phonetic inventory when sampled in conversational speech. The statistically significant differences in PCI Total scores for the Total group comparisons, as well as in the PCI Middle-8 and PCI Total for the OME severity subgroup comparisons, indicate that Study 2 children in the OME+ group had less well-developed consonant inventories than children with OME– histories.

4. *Vowel/diphthong errors.* The significant between-group difference on the Total group PVC comparison indicates that Study 2 children with OME+ histories had lower vowel/diphthong accuracy in conversational speech than children with OME– histories.

5. *Vowel/diphthong substitution errors.* The significant between-group differences on the PVC-R in both the Total group and OME severity subgroup comparisons indicate that Study 2 children with OME+ histories made significantly more vowel/diphthong substitution errors (for technical reasons, vowel deletion errors are excluded from this analysis).

Note that the statistically significant findings described in 1, 2, 4, and 5 above are reflected also in findings for the PPC and PPC-R comparisons, which index articulation competence on both consonants and vowels/diphthongs. As shown in Table 3, Study 2 children with OME+ histories had significantly lower PPC scores in the Total group comparison and significantly lower scores on the PPC-R on both the Total group and OME severity subgroup comparisons.

6. *Intelligibility.* Finally, as shown in the bottom row of Table 3, Intelligibility Index percentages were significantly different for both the Total group and the OME severity subgroup comparisons. Compared to children with OME– histories, Study 2 children with OME+ histories had significantly fewer words that were intelligible to the examiner and/or transcriber.

Individual Risk Analyses

The preceding analyses assessed group-level differences on speech measures treated as continuous variables. An alternative approach used in epidemiologic and other public health research is to treat disorder as a qualitative variable, using criteria to classify each individual as affected or not affected. As described previously, the 10th speech measure, the Speech Disorders Classification System (SDCS), classifies a child as normal (i.e., Normal [or Normalized] Speech Acquisition [NSA]), subclinical disorder (Normal [or Normalized] Speech Acquisition/Speech Delay [NSA/SD]), or clinical disorder (Speech Delay [SD]). Procedures were also needed to classify children as normal, subclinical, or clinical on each of the other nine speech metrics. The following three sections describe how this was accomplished, the procedures used for risk analyses, and an estimate of the validity of the classification methods.

Classification of Children on the Nine Speech Measures

Classification of children on each of the nine speech measures and their subscales was accomplished using lifespan reference data (means, standard deviations) for each metric for boys and girls at each age from 3 to 8 years (Austin & Shriberg, 1996). The reference database, including files for 321 children in this age range (170 boys, 151 girls), was assembled from conversational speech samples from several Midwestern cities, with all children meeting criteria for normal or normalized speech acquisition on the SDCS (Shriberg, 1993; Shriberg et al., 1997b).

The first step in the classification procedure was to calculate z scores for each of the nine speech measures for all participants in Study 1 and Study 2 using the appropriate Age × Gender means and standard deviations. The second step was to determine the cutoff criteria for z scores to sort all children into three groups: *normal speech*, *subclinical disorder*, and *clinical disorder*. Three considerations influencing the cutoff criteria were (a) the magnitudes of the standard errors of measurement for each of the nine metrics (Shriberg et al., 1997a); (b) rationale proposed in several epidemiologic (Tomblin, Records, & Zhang, 1996), genetics (Lahey & Edwards, 1995), and comorbidity (Shriberg & Austin, 1998) studies addressing z score cutoff criteria for Specific Language Impairment (SLI); and (c) an examination of the distribution of z scores in the present data. As in each of the studies just cited, definition of clinical and subclinical speech disorder for the Total group analyses in the present study used more stringent cutoff criteria than the customary criterion of one standard deviation or greater below the mean performance for each gender by age. The primary rationale for setting more

stringent cutoff criteria was that the reference data reflected the range of performance of children whose speech was classified as normal (NSA) by the 10th speech measure, the SDCS. Thus, the standard deviations in the reference data reflected the variability of normal scores, not the variability of scores in a sample of a population including persons with the disorder. Accordingly, the criterion for subclinical disorder for the Total group analyses was a *z* score from –1.3 to –1.9. That is, scores of 1.3 to 1.9 standard deviation units below the mean of children with normal speech are approximately equivalent to the 9.7 to 2.3 percentiles of the reference group. The criterion for clinical disorder for the Total group analyses was a *z* score of –2.0 or below, which corresponds to scores at or below the 2.3 percentile of the reference data of children with normal speech.

Relative Risk Ratio and Odds Ratio Analyses

Relative risk ratio and odds ratio analyses are two model-free methods that are well suited to the task of assessing the risk of clinical or subclinical speech disorder in children with histories of OME (cf. Kahn & Sempos, 1989; Khoury, Beaty, & Cohen, 1993). It is important to describe the conceptual and operational differences between the two ratios, both of which are used in the following risk estimate analyses.

Relative risk ratios express the risk of a disease or disorder for persons exposed to a specified risk factor relative to the risk of a disease in unexposed persons. In the present data, the sampling and selection procedures used in Study 1 and for the original subsample of 28 children in Study 2 qualify for relative risk analyses because each of the children in these cohorts was classified by exposure history (i.e., OME+ or OME–). The relative risk ratio thus expresses the percentage of children exposed to the risk factor (OME+) who are affected with a disease (speech disorder) compared to the percentage of nonexposed, affected children. A relative risk of 1 reflects no difference in percentages, whereas departures from 1 indicate decreased (i.e., protective) or increased risk for disorder with exposure. In the present study, 95% confidence limits were used to estimate the lower and upper boundaries for the relative risk ratios (and odds ratios; see below). Confidence intervals represent the range within which the true magnitude of effect lies with a certain degree of assurance. The width of the confidence interval indicates the amount of variability inherent in the estimate of risk and thus the effect of sample size.

Odds ratios are an alternative to relative risk ratios in case-control designs or cross-sectional studies in which the participants do not constitute one true cohort. Odds ratios also express the risk associated with exposure to disease, but they do not have the predictive power associated with relative risk ratios because they are not based on outcomes for a cohort. In the present study, odds ratios were used as estimates of the risk for speech disorder associated with OME histories for the 18 children constituting the OME severity subgroup in Study 1 and the 50 children making up the combined Total group for Study 2.

Validity of the Classification Methods

Rationale. Table 4 is a summary of the risk ratio analyses for children in the Total groups and OME severity subgroups in Study 1 and Study 2. A validity estimate for the speech classification procedures can be obtained by inspecting the percentages of OME– children in each study (i.e., the "controls") who are classified as having clinical involvement on each of the nine speech measures. Prevalence estimates for speech disorder in preschool children can be used to set the expected liability (prevalence rate) for speech disorder in the groups of children with OME– histories. However, there are several sociodemographic characteristics of the present samples that might be associated with a higher-than-expected prevalence of clinical disorder in the two OME– groups. First, a higher prevalence of clinical disorder might be expected in Study 1 because all of the children are 3 years old at the lowest end of the preschool age range. That is, prevalence should be higher in the youngest age group because a child with speech disorder had little opportunity to normalize. Second, higher prevalence rates for speech disorder also might be expected for the OME– children in Study 2 because of sociodemographic factors. As described previously, all OME+ and OME– children in Study 2 were attending a Head Start program on the reservation. Finally, higher prevalence rates in both studies might be expected because the present study used conversational speech and more fine-grained transcription than the articulation tests and broad transcription used in available prevalence studies. It is not possible to assess the validity of the cutoff points for subclinical involvement because there are no prevalence estimates that are conceptually similar to this view of disorder as a semicontinuous trait (i.e., as indicated above by the classification NSA/SD; cf. Shriberg, 1993).

Findings. Support for the validity of the *z* score cutoff criteria used to classify a child as affected on each of the nine speech metrics, their subscales, and the SDCS is based on comparison of the resulting percentages to the unconditional expected prevalence of speech disorder, plus the four considerations reviewed above. The most widely cited estimate of the prevalence of clinical

Table 4. Risk estimates (relative risk ratio and odds ratio analyses) for children in Study 1 and Study 2. Outcome variables are dichotomized as normal versus subclinical and clinical.[a]

		Total group						OME severity subgroup					
		% Affected		Estimate		Confidence limits		% Affected		Estimate		Confidence limits	
		OME+	OME–	Risk	Power	Lower	Upper	OME+	OME–	Risk	Power	Lower	Upper
Study 1													
PCC	Early	5	19	0.28	0.06	0.03	2.44	0	22	0.00	0.24	0.00	5.22
	Middle	11	6	1.68	0.06	0.17	16.91	0	0	1.00[c]	—[d]	0.02	55.80
	Late	0	6	0.00	0.11	0.00	32.84	0	0	1.00[c]	—[d]	0.02	55.80
	Total	5	0	2.68[c]	0.09	0.10	70.31	0	0	1.00[c]	—[d]	0.02	55.80
PCC-A	Middle	11	6	1.68	0.06	0.17	16.91	0	0	1.00[c]	—[d]	0.02	55.80
	Late	21	13	1.68	0.10	0.35	8.03	22	11	2.29	0.09	0.09	151.42
	Total	16	19	0.84	0.04	0.20	3.61	0	22	0.00	0.24	0.00	5.22
PCC-R	Early	5	13	0.42	0.12	0.04	4.23	0	11	0.00	0.11	0.00	39.00
	Middle	21	6	3.37	0.06	0.42	27.18	0	0	1.00[c]	—[d]	0.02	55.80
	Late	16	13	1.26	0.05	0.24	6.65	11	11	1.00	—[d]	0.01	88.19
	Total	21	19	1.12	0.04	0.29	4.29	0	22	0.00	0.24	0.00	5.22
PCI	Early	0	0	0.90[c]	0.03	0.02	50.25	0	0	2.60[c]	0.07	0.04	170.39
	Middle	42	19	2.25	0.17	0.71	7.08	11	22	0.44	0.09	0.01	10.63
	Late	21	31	0.67	0.10	0.22	2.09	11	33	0.25	0.20	0.00	4.37
	Total	32	31	1.01	0.03	0.38	2.70	11	33	0.25	0.20	0.00	4.37
PVC		21	25	0.84	0.05	0.25	2.84	0	33	0.00	0.39	0.00	2.23
PVC-R		16	13	1.26	0.05	0.24	6.65	11	22	0.44	0.09	0.01	10.63
PPC		5	13	0.42	0.12	0.04	4.23	0	22	0.00	0.24	0.00	5.22
PPC-R		16	19	0.84	0.04	0.20	3.61	0	22	0.00	0.24	0.00	5.22
Int. Index		58	38	1.54	0.10	0.74	3.23	67	33	4.00	0.29	0.41	43.35
SDCS[b]		37	6	5.89	0.37	0.81	42.99	0	0	1.00[c]	—[d]	0.02	55.80
Study 2													
PCC	Early	37	13	3.94	0.51	0.79	21.42	50	21	2.33	0.18	0.75	7.23
	Middle	26	13	2.41	0.22	0.43	13.99	36	14	2.50	0.10	0.58	10.80
	Late	21	6	3.87	0.32	0.48	46.03	29	14	2.00	0.15	0.43	9.21
	Total	26	3	10.71	0.66	1.01	523.12*	36	7	5.00	0.24	0.67	37.51
PCC-A	Middle	32	16	2.40	0.25	0.49	11.83	43	21	2.00	0.09	0.62	6.45
	Late	37	13	3.94	0.51	0.79	21.42	43	14	3.00	0.20	0.73	12.39
	Total	53	16	5.78	0.78	1.31	26.96*	64	21	3.00	0.43	1.02	8.80*
PCC-R	Early	37	10	5.44	0.65	0.99	36.83	50	14	3.50	0.31	0.88	13.99
	Middle	42	26	2.09	0.22	0.52	8.32	50	29	1.75	0.10	0.66	4.66
	Late	47	16	4.68	0.66	1.06	21.88*	57	14	4.00	0.46	1.03	15.60*
	Total	53	19	4.63	0.69	1.10	20.02*	64	29	2.25	0.29	0.90	5.62
PCI	Early	0	0	1.84[c]	0.05	0.03	100.45	0	0	1.00[c]	—[d]	0.02	56.47
	Middle	11	13	0.79	0.04	0.07	6.29	14	7	2.00	0.09	0.20	19.62
	Late	32	10	4.31	0.51	0.75	29.90	36	14	2.50	0.10	0.58	10.80
	Total	26	13	2.41	0.22	0.43	13.99	29	14	2.00	0.15	0.43	9.21
PVC		21	0	18.29[c]	0.70	0.92	361.69	29	0	12.43[c]	0.50	0.60	256.68
PVC-R		16	0	13.36[c]	0.40	0.65	274.53	21	0	8.83[c]	0.36	0.41	188.74
PPC		26	3	10.71	0.66	1.01	523.12*	36	7	5.00	0.24	0.67	37.51
PPC-R		53	13	7.50	0.86	1.59	39.52*	64	21	3.00	0.43	1.02	8.80*
Int. Index		16	10	1.75	0.10	0.21	14.51	21	0	8.83[c]	0.36	0.41	188.74
SDCS[b]		32	10	4.31	0.51	0.75	29.90	36	7	5.00	0.24	0.67	37.51

[a]Table 2 provides *n*s for the OME+ and OME– subgroups for Study 1 and Study 2. Risk is estimated by relative risk (RR) for Study 1: Total group and Study 2: OME severity subgroup, and odds ratio (OR) for Study 1: OME severity subgroup and Study 2: Total group. [b]SDCS = Speech Disorders Classification System. NSA vs. NSA/SD and SD. See Table 3 for other measure abbreviations. [c]RR and OR cannot be calculated when none of the OME– children is affected. In those cases, risk is estimated by adding 0.5 to the counts for each cell and calculating the odds ratio. Confidence limits are estimated as

$$\exp\left(\ln OR \pm 1.96\sqrt{\frac{1}{a+.5}+\frac{1}{b+.5}+\frac{1}{c+.5}+\frac{1}{d+.5}}\right)$$

(Kahn & Sempos, 1989). [d]Power cannot be calculated.

*risk significantly greater than 1.

speech disorder in preschool children is 3–5% (Leske, 1981; Winitz & Darley, 1980), but a recent prevalence study of 1,328 6-year-old children estimates its prevalence at that older age at 3.8% (Shriberg, Tomblin, & McSweeny, 1998). Using preliminary normalization data, Shriberg, Tomblin et al. (1998) estimate that prevalence of speech delay at exactly 3 years of age may be as high as approximately 14%. For the OME– children in the present two studies, the prevalence of clinical or subclinical disorder ranged from 0% to 38%. As shown in Table 4, 28 of the 42 (67%) estimates for the Total groups were 13% or lower. Given that both subclinical and clinical disorder are included in these estimates, together with the four moderating factors that would increase the prevalence of speech disorder in these children, these data are interpreted as providing strong validity support for the cutoff score criteria used to classify children for the risk estimate analyses.

Risk Estimate Findings

The findings in Table 4 provide information on the increased risk for subclinical or clinical speech disorder in children with OME+ histories. All analyses were completed using the STATCALC module in the Epi Info program (Dean et al., 1995), which provided risk estimates and 95% confidence limits for the risk estimates. Confidence intervals not including 1 are statistically significant. Power estimates calculated as previously described are also included for each analysis.

Beginning with the data for children in Study 1, there were no statistically significant findings indicating increased risk for speech disorder in children with OME+ histories in either the Total group or OME severity subgroup analyses.

In Study 2, the Total group analyses yielded nine statistically significant increases in the risk for subclinical or clinical speech disorder in children with OME+ histories. These findings were obtained on the total and/or subscales of 5 of the 10 different speech metrics. Significantly increased risk for lowered performance on these five metrics for children with OME+ histories ranged from 3 times the risk of children in the OME– groups (OME severity subgroup analyses: PCA-Total: CI = 1.02–8.80 and PPC-R: CI = 1.02–8.80) to 10.71 increased risk (Total group analyses: PCC-Total: CI = 1.01–523.12 and PPC: CI = 1.01–523.12). As indicated by the width of the confidence level and the power estimate, the precision and stability of some of these estimates is not high. The percentage of children with OME+ histories classified as having subclinical or clinical disorder (i.e., "affected") in these statistically significant comparisons ranged from 14% (PCC-R Late-8 consonants) to 53% (PCC-A Total, PCC-R Total, and PPC-R).

Discussion

Several perspectives on the between-group and individual risk analysis findings support the conclusion that a significant history of early recurrent otitis media with effusion placed the children in Study 2 at increased risk for speech disorder.

Measurement

From a measurement perspective, it is significant that OME history was associated with differences on both the Intelligibility Index and the PCC-R. The Intelligibility Index has inherent face validity as a metric reflecting speech disability because it includes contributions from speech, language, and prosodic variables (Shriberg & Kwiatkowski, 1982). As well, the PCC-R has been proposed as the most theoretically sensitive and psychometrically stable index of speech-sound production of the nine speech measures (Shriberg et al., 1997a). In eliminating the conceptual and reliability issues associated with speech-sound distortions, the PCC-R directly reflects the percentage of consonant deletion and substitution errors in conversational speech. Significant between-group differences were found on the PCC-R for both the groupwise OME+ versus OME– comparisons and for the risk estimate comparisons. For the latter, 53% of the Study 2 children from Native American backgrounds with OME+ histories met criteria for a subclinical or clinical disorder compared to 19% of children with OME– histories—a 34% difference. The odds ratio for these findings indicates a 4.63 increased risk for speech disorder, with a 95% confidence limit of 1.10 to 20.2 bounding this ratio. The power to detect this difference was estimated at .69. Given the small standard error of measurement for the PCC-R (estimated at 2.4 percentage points for boys and girls in this age range; Shriberg et al., 1997a), these findings are secure relative to potential measurement error. However, the width of the confidence interval indicates the imprecision of the exact risk, likely due to the small sample sizes. These statistically significant findings on the Intelligibility Index and the PCC-R are viewed as providing the best criterion validity for the conclusion that a significant history of OME increased the risk for speech disorder in the children in Study 2.

Participants

Although descriptive trends in both Study 1 and Study 2 indicated poorer speech acquisition in the OME+ compared to OME– groups, the statistically significant findings were nearly all limited to Study 2 comparisons. One obvious between-group difference was the reduced statistical power in Study 1 compared to Study 2. However,

there were also major differences in the sociodemographic composition of children in these two groups involving age, race/ethnicity, and social class. Moreover, there were many potentially confounding subject variables that were unmeasured in these data (e.g., gestational age, birth weight, familial history of speech disorder). On differences in age, one suggestion in the otitis literature is that 3 years of age (Study 1) is too early to draw conclusions about long-term effects of middle-ear disease and its correlates on speech (cf. Paul, Lynn, & Lohr-Flanders, 1993; Roberts et al., 1988). Clinically, however, children whose speech is delayed at 3 years of age (or even at younger ages) are routinely referred for early intervention. For the present research concern, note that the speech error patterns of younger children should be more sensitive to the potential effects of OME than is the case with older children because of younger children's reduced opportunity for speech-sound normalization. On health care issues, the medical records of children in both studies indicated regular and excellent health care. Early and appropriate attention by physicians and caregivers has been cited as one of the major preventive variables in studies using diverse subject pools and independent and outcome measures (e.g., Apel & Marazzi, 1994; Black & Sonnenschein, 1993; Feldman & Gelman, 1986; Freeark et al., 1992; Paden, Matthies, & Novak, 1989; Roberts et al., 1995).

Thus, the contrast in the number of significant findings in the two studies, each of which used identical speech measures and methods, demonstrates the crucial role of subject characteristics in OME-speech research. The young participants in Study 1 were generally closest in sociodemographic characteristics to children in many of the otitis-speech studies conducted in communities where children have received good health care and are from language-rich learning environments (cf. Table 1). Moreover, at 3 years of age, the OME+ and OME– children had less chance to normalize speech errors than have older children in OME-speech studies. In contrast, the somewhat older participants in Study 2 are more comparable to otitis-speech studies conducted with children who may not be experiencing those early home and environmental factors that mitigate the effects of the chronic discomfort and hearing loss associated with early recurrent OME. Related issues, as they affect language acquisition, are addressed in the second paper in this series. As discussed next, the contrast in findings, using identical speech assessment methods, underscores the need for multifactorial models relating early recurrent OME to later speech disorder.

Effect Size

The present findings are consistent with prior literature indicating small to moderate effect sizes, both in the number of children with OME who have a speech disorder and the severity of their speech involvement. Small effect sizes are especially apparent in the severity differences as shown for the grouped data in Table 3. When divided into normal, subclinical, and clinical involvement, however, risk trends indicate more subclinical or clinical involvement for children with OME+ histories, although not all were statistically significant. Most telling on this issue are the SDCS findings, for which there were few statistically significant differences associated with OME histories. As described elsewhere (Shriberg et al., 1997b), the stringent developmental criteria for speech delay on the SDCS often result in some children clinically referred for delayed speech being classified as either subclinical delay or normal(ized). To the extent that other studies listed in Table 1 have used measures of comparable stringency, they may not have been sensitive to the differences identified by the other speech metrics used in the present study.

As suggested previously, the relatively small effect sizes might also serve to document the large roles other organismic and environmental variables may play in mitigation or prevention. Most otitis-speech researchers endorse diverse multifactorial explanatory models as explanations for small effect sizes (e.g., Abraham, Wallace, & Gravel, 1996; Bishop & Edmundson, 1986; Paradise, 1997; Paul et al., 1993; Peters, Grievink, van Bon, van den Bercken, & Schilder, 1997; Ruben, 1984; Shriberg, 1987; Vernon-Feagans, 1997). Hall and Hill (1986), for example, discussed the need to consider variables in four explanatory domains: otitis media, hearing loss, the child, and the environment. Although there were many differences in the child and environment characteristics of children in Study 1 and Study 2, both groups received a high level of medical care, which should have minimized both the physical discomfort of OME and associated hearing loss. Thus, other than age differences in the two study samples as possible explanatory sources, child and environment factors that were not assessed in this study likely moderated or mediated the primary findings.

The precursor variable notably missing from these data and likely to have the strongest association with effect size is information on children's hearing levels during OME episodes. The few data associating the amount and period of measured or inferred hearing loss with later speech disorder are equivocal. For example, Hubbard, Paradise, McWilliams, Elster, and Taylor (1985) attributed the difference in eventual articulation competence observed in two groups of adolescent children with clefts of the palate to the positive effects on hearing levels of early insertion of pressure equalization tubes, with the more articulate speakers receiving tubes on average at 3.0 months and the other group on average at 30.8 months. However, Paden et al. (1989) found that elapsed time of OME until remission was a

predictor of speech normalization, but hearing loss was not associated with specific types of error patterns. In commenting on the potential sources of variance in their data, Paden and colleagues also stressed the need for multifactorial causal models that include child and environment variables:

> The most important finding of this study...is that phonological delay at age 3 in children with histories of frequently recurring OME cannot be attributed to a single factor. Children cannot be identified as at risk simply by calculating the amount of time during which OME has been experienced, or by assessing the severity of the hearing impairment it has caused, or even by a combination of these measures. (p. 240)

Conclusion

The central finding of this report is that early recurrent otitis media with effusion was associated with approximately 4.6 (CI = 1.10–20.02) increased risk for subclinical or clinical speech disorder in a demographically well-controlled sample of 50 children and was not associated with significantly increased risk for a different demographically well-controlled sample of 35 children. Because OME measures were comparable in the two samples and speech measures were identical, these data are interpreted as supporting a multifactorial model relating early recurrent OME to later speech disorder. In addition to otological, audiological, and age variables that may have been associated with the differences in outcomes, diverse child and environment factors may have contributed to the risk for speech disorder in one study sample and mitigated risk in the other. Implications for intervention and prevention must await findings from large and well-controlled prospective studies that allow for multifactorial modeling of child and environmental variables and that explicate the psycholinguistic processes linking early recurrent OME to later communicative styles and impairments.

Acknowledgments

Preparation of this report was supported by the National Institute on Deafness and Other Communication Disorders, National Institutes of Health (DC00496). Many people contributed their time and expertise to these studies. We extend our sincere thanks to Chad Allen, Diane Austin, Delores Boyd, Betty Bricco, Carol Dodge, James Goinz, Frederic Gruber, Rebecca Hinke, Barbara Hodson, Kit Hoffmann, Patricia Engebose Hovel, Maureen McGowan Jepsen, Peter Karofsky, Roger Klumb, Gregory Lof, Jane McSweeny, Amparo Ortez, Mary Matz, Mary Jo Osberger, Elaine Paden, Mary Beth Pecore, Carmen Rasmussen, Dorothy Rorick Ross, Anne Smith, Geralyn Timmer, Catherine Trost-Steffen, Sandy Wallin, Jerry Waukau, Cindy Weisflock, Audrey Weston, Carol Widder, Terry Wiley, and David Wilson. We also thank the editor, associate editor, and four reviewers for their extremely valuable editorial critique. Copies of relevant technical reports are available at the Phonology Project web site: http://www.waisman.wisc.edu/phonology/.

References

Abraham, S. S., Wallace, I. F., & Gravel, J. S. (1996). Early otitis media and phonological development at age 2 years. *Laryngoscope, 106*, 727–732.

Al-Sunduqchi, M. S. (1990). *Determining the appropriate sample size for inferences based on the Wilcoxon Statistics*. Unpublished doctoral dissertation, University of Wyoming, Laramie.

ANSI. (1969). *Specifications for audiometers* (ANSI S3.1969). New York: Author.

Apel, K., & Marazzi, S. (1994, November). *Effects of otitis media on lexical learning*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA.

Austin, D., & Shriberg, L. D. (1996). *Lifespan reference data for ten measures of articulation competence using the Speech Disorders Classification System (SDCS)* (Tech. Rep. No. 3). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Bess, F. H. (1983). Hearing loss associated with middle ear effusion. *Pediatrics, 71*, 640–641.

Bishop, D. V. M., & Edmundson, A. (1986). Is otitis media a major cause of specific developmental language disorders? *British Journal of Disorders of Communication, 21*, 321–338.

Black, M. M., & Sonnenschein, S. (1993). Early exposure to otitis media: A preliminary investigation of behavioral outcome. *Developmental and Behavioral Pediatrics, 14*(3), 150–155.

Churchill, J. D., Hodson, B. W., Jones, B. W., & Novak, R. E. (1988). Phonological systems of speech-disordered clients with positive/negative histories of otitis media. *Language, Speech, and Hearing Services in Schools, 19*, 100–107.

Daly, K. A. (1997). Definition and epidemiology of otitis media. In J. E. Roberts, I. F. Wallace, & F. W. Henderson, *Otitis media in young children: Medical, developmental, and educational considerations* (pp. 3–41). Baltimore, MD: Paul H. Brookes Publishing Co.

Dean, A. G., Dean, J. A., Coulombier, D., Brendel, K. A., Smith, D. C., Burton, A. H., Dicker, R. C., Sullivan, K., Fagan, R. F., & Arner, T. G. (1995). Epi Info, Version 6: a Word-processing, Database, and Statistics Program for Public Health on IBM-compatible Microcomputers [Computer software]. Atlanta, GA:

Centers for Disease Control and Prevention.

Dobie, R. A., & Berlin, C. I. (1979). Influence of otitis media on hearing and development. *Annals of Otology, Rhinology and Laryngology, 88* (part 2, supplement 60), 48–53.

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

Dyson, A. T., Holmes, A. E., & Duffitt, D. V. (1987). Speech characteristics of children after otitis media. *Journal of Pediatric Health Care, 1*, 261–265.

Elliott, L. L., & Katz, D. R. (1980). *Northwestern University Children's Perception of Speech (NU-CHIPS).* St. Louis, MO: Auditec.

Feagans, L., Sanyal, M., Henderson, F., Collier, A., & Applebaum, M. (1987). Relationship of middle ear disease in early childhood to later narrative and attentional skills. *Journal of Pediatric Psychology, 12*, 581–594.

Feldman, H., & Gelman, R. (1986). Otitis media and cognitive development: Theoretical perspectives. In J. F. Kavanaugh (Ed.), *Otitis media and child development* (pp. 27–41). Parkton, MD: York.

Freeark, K., Frank, S. J., Wagner, A. E., Lopez, M., Olmsted, C., & Girard, R. (1992). Otitis media, language development, and parental verbal stimulation. *Journal of Pediatric Psychology, 17*(2), 173–185.

Goinz, J. (1984). Otitis media among preschool and school-age Indian children in MI, MN and WI. *Hearing Instruments, 35*(6), 16–18.

Hall, D. M. B., & Hill, P. (1986). When does secretory otitis media affect language development? *Archives of Disease in Childhood, 61*, 42–47.

Harsten, G., Nettelbladt, U., Schalen, L., Kalm, O., & Prellner, K. (1993). Language development in children with recurrent otitis media during the first three years of life: Follow-up study from birth to seven years of age. *Journal of Laryngology and Otology, 107*, 407–412.

Hintze, J. L. (1996). Power Analysis and Sample Size (Version 6.0) [Computer software]. Kaysville, UT: NCSS.

Hoey Hemmer, V., & Bernstein Ratner, N. (1994). Communicative development in twins with discordant histories of otitis media. *Journal of Communication Disorders, 27*, 91–106.

Holm, V. A., & Kunze, L. H. (1969). Effect of chronic otitis media on language and speech development. *Pediatrics, 43*, 833–839.

Hubbard, T. W., Paradise, J. L., McWilliams, B. J., Elster, B. A., & Taylor, F. H. (1985). Consequences of unremitting middle-ear disease in early life: Otologic, audiologic, and developmental findings in children with cleft palate. *The New England Journal of Medicine, 312*, 1529–1534.

Kahn, H. A., & Sempos, C. T. (1989). *Statistical methods in epidemiology*. New York: Oxford University Press.

Kavanagh, J. F. (Ed). (1986). *Otitis media and child development*. Parkton, MD: York Press.

Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Kertoy, M. K., & Shriberg, L. D. (1984, November). *Language correlates of recurrent otitis media*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, San Francisco, CA.

Khoury, M. J., Beaty, T. H., & Cohen, B. H. (1993). *Fundamentals of genetic epidemiology.* New York: Oxford University Press.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities (rev. ed.).* Urbana, IL: University of Illinois Press.

Kwiatkowski, J., & Shriberg, L. D. (1992). Intelligibility assessment in developmental phonological disorders: Accuracy of caregiver gloss. *Journal of Speech and Hearing Research, 35*, 1095–1104.

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.

Lehmann, M. D., Charron, K., Kummer, A., & Keith, R. W. (1979). The effects of chronic middle ear effusion on speech and language development. *International Journal of Pediatric Otorhinolaryngology, 1*, 137–144.

Leske, M. (1981). Prevalence estimates of communicative disorders in the United States: Speech disorders. *Asha, 23*, 217–228.

Lonigan, C. J., Fischel, J. E., Whitehurst, G. J., Arnold, D. S., & Valdez-Menchaca, M. C. (1992). The role of otitis media in the development of expressive language disorder. *Developmental Psychology, 28*, 430–440.

Lous, J. (1990). Secretory otitis media and phonology when starting school. *Scandinavian Audiology, 19*, 215–222.

Lous, J. (1995). Secretory otitis media in schoolchildren. *Danish Medical Bulletin, 42*(1), 71–99.

Manders, E., & Tyberghein, J. (1993). The effects of ventilation tube placement on hearing, speech, language, cognition and behavior. *Acta Oto-Rhino-Laryngologica Belgica, 47*, 27–32.

McReynolds, L., & Elbert, M. (1981). Criteria for phonological analysis. *Journal of Speech and Hearing Disorders, 46*, 197–203.

McSweeny, 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.

Miller, J. F., & Yoder, D. E. (1984). *Miller-Yoder Language Comprehension Test*. Austin, TX: Pro-Ed.

Mustain, W. D. (1979). Linguistic and educational implications of recurrent otitis media. *Ear, Nose and Throat Journal, 58*, 218–222.

Needleman, H. (1977). Effects of hearing loss from early recurrent otitis media on speech and language development. In B. Jaffe (Ed.), *Hearing loss in children* (pp. 640–649). Baltimore, MD: University Park Press.

Nozza, R. J. (1988). Auditory deficit in infants with otitis media with effusion: More than a "mild" hearing loss. In D. J. Lim, C. D. Bluestone, J. O. Klein, & J. D. Nelson (Eds.), *Recent advances in otitis media* (pp. 376–379). Philadelphia: B. C. Decker, Inc.

Paden, E. P., Matthies, M. L., & Novak, M. A. (1989). Recovery from OME-related phonologic delay following tube placement. *Journal of Speech and Hearing Disorders, 54*, 94–100.

Paden, E. P., Novak, M. A., & Beiter, A. L. (1987). Predictors of phonologic inadequacy in young children prone to otitis media. *Journal of Speech and Hearing Disorders, 52*, 232–242.

Paradise, J. L. (1997). Developmental outcomes in relation to early-life otitis media: Present and future directions in research. In J. E. Roberts, I. F. Wallace, & F. W. Henderson, *Otitis media in young children: Medical, developmental, and educational considerations* (pp. 287–306). Baltimore, MD: Paul H. Brookes Publishing Co.

Paul, R., Lynn, T. F., & Lohr-Flanders, M. (1993). History of middle ear involvement and speech/language development in late talkers. *Journal of Speech and Hearing Research, 36*, 1055–1062.

Pendergast, K., Dickey, S., Selmar, J., & Soder, A. (1969). *The Photo Articulation Test*. Danville, IL: Interstate Printers and Publishers.

Peters, S. A. F., Grievink, E. H., van Bon, W. H. J., van den Bercken, J. H. L., & Schilder, A. G. M. (1997). The contribution of risk factors to the effect of early otitis media with effusion on later language, reading, and spelling. *Developmental Medicine & Child Neurology, 39*, 31–39.

Rabbitt, P. M. A. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology, 20*, 241–248.

Raymond, C., Garcia, P., & Scheib, R. (1993, November). *Prevalence of hearing disorders in Native American preschoolers*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Anaheim, CA.

Roberts, J. E., Burchinal, M. R., & Campbell, F. (1994). Otitis media in early childhood and patterns of intellectual development and later academic performance. *Journal of Pediatric Psychology, 19*(3), 347–367.

Roberts, J. E., Burchinal, M. R., Davis, B. P., Collier, A. M., & Henderson, F. W. (1991). Otitis media in early childhood and later language. *Journal of Speech and Hearing Research, 34*, 1158–1168.

Roberts, J. E., Burchinal, M. R., Koch, M. A., Footo, M. M., & Henderson, F. W. (1988). Otitis media in early childhood and its relationship to later phonological development. *Journal of Speech and Hearing Disorders, 53*, 424–432.

Roberts, J. E., & Clarke-Klein, S. (1994). Otitis media. In J. E. Bernthal & N. W. Bankson (Eds.), *Child phonology: Characteristics, assessment, and intervention with special populations* (pp. 182–198). New York: Thieme Medical Publishers, Inc.

Roberts, J. E., Gravel, J., Schwartz, R. G., Dollaghan, C., & Campbell, T. (1998, November). *Does otitis media cause speech and language sequelae?* Paper presented at Annual Convention of the American Speech-Language-Hearing Association, San Antonio, TX.

Roberts, J. E., Wallace, I. F., & Henderson, F. W. (Eds.). (1997). *Otitis media in young children: Medical, developmental, and educational considerations*. Baltimore, MD: Paul H. Brookes Publishing Co.

Roberts, K. (1997). A preliminary account of the effect of otitis media on 15-month-olds' categorization and some implications for early language learning. *Journal of Speech, Language, and Hearing Research, 40*, 508–518.

Ruben, R. J. (1984). An inquiry into the minimal amount of auditory deprivation which results in a cognitive effect in man. *Acta Otolaryngol (Stockh), 414*, 157–164.

Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L. (1996a). Formant frequencies of vowels produced by infants with and without early onset otitis media. *Canadian Acoustics, 24*(2), 19–28.

Rvachew, S., Slawinski, E. B., Williams, M., & Green, C. L. (1996b). The impact of early onset otitis media on prelinguistic speech development. *Journal of Speech-Language Pathology and Audiology, 20*, 247–255.

Schlieper, A., Kisilevsky, H., Mattingly, S., & Yorke, L. (1985). Mild conductive hearing loss and language development: A one-year follow-up study. *Developmental and Behavioral Pediatrics, 6*(2), 65–68.

Shriberg, L. D. (1986). *PEPPER: Programs to Examine Phonetic and Phonologic Evaluation Records*. Hillsdale, NJ: Lawrence Erlbaum.

Shriberg, L. D. (1987). In search of the otitis media-speech connection. *Journal of the National Student Speech Language Hearing Association, 15*, 56–67.

Shriberg, L. D. (1993). 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., & 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 Publishing Co.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, 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., McSweeny, 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., Friel-Patti, S., Flipsen, P., Jr., & Brown, R. L. (2000). Otitis media, fluctuant hearing loss, and speech-language delay: A preliminary structural equation model. *Journal of Speech, Language, and Hearing Research, 43*, 100–120.

Shriberg, L. D., & Kent, R. D. (1982). *Clinical phonetics*. New York: Macmillan.

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 III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47*, 256–270.

Shriberg, L. D., Kwiatkowski, J., Block, M. G., Katcher, M. L., Kertoy, M. K., & Nellis, R. A. (1984, November).

Phonetic and phonologic correlates of early recurrent otitis media. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, San Francisco, CA.

Shriberg, L. D., Kwiatkowski, J., & Hoffmann, K. A. (1984). A procedure for phonetic transcription by consensus. *Journal of Speech and Hearing Research, 27,* 456–465.

Shriberg, L. D., & Smith, A. J. (1983). Phonological correlates of middle-ear involvement in speech-delayed children: A methodological note. *Journal of Speech and Hearing Research, 26,* 293–297.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1998, November). *Prevalence of speech delay in kindergarten-aged children*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, San Antonio, TX.

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Silva, P. A., Chalmers, D., & Stewart, I. (1986). Some audiological, psychological, educational and behavioral characteristics of children with bilateral otitis media with effusion: A longitudinal study. *Journal of Learning Disabilities, 19,* 165–169.

Silva, P. A., Kirkland, C., Simpson, A., Stewart, I. A., & Williams, S. M. (1982). Some developmental and behavioral problems associated with bilateral otitis media with effusion. *Journal of Learning Disabilities, 15,* 417–421.

Stewart, J. (1986). Hearing disorders among the indigenous peoples of North America and the Pacific Basin. In O. Taylor (Ed.), *Nature of communication disorders in culturally and linguistically diverse populations* (pp. 237–276). San Diego, CA: College-Hill Press.

Teele, D. W., Klein, J. O., Chase, C., Menyuk, P., & Rosner, B. A. (1990). Otitis media in infancy and intellectual ability, school achievement, speech, and language at age 7 years. *Journal of Infectious Diseases, 162,* 685–694.

Teele, D. W., Klein, J. O., Rosner, B. A., & The Greater Boston Otitis Media Study Group. (1984). Otitis media with effusion during the first three years of life and development of speech and language. *Pediatrics, 74*(2), 282–287.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

Thielke, H. M. (1988). *Speech and language sequelae of recurrent otitis media in Menominee Indian preschool children*. Unpublished in-lieu-of master's thesis, University of Wisconsin, Madison.

Thielke, H. M. , & Shriberg, L. D. (1990). Effects of recurrent otitis media on language, speech, and educational achievement in Menominee Indian children. *Journal of American Indian Education, 29,* 25–35.

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.

Toubbeh, J. I. (1985). *Otitis media in the American Indian and Alaskan Native populations.* Unpublished manuscript.

Trammel, J. (1977). *Test of Auditory Comprehension*. North Hollywood, CA: Foreworks.

van der Vyver, M., van der Merwe, A., & Tesner, H. E. C. (1988). The effects of otitis media on articulation in children with cerebral palsy. *International Journal of Behavioral Research, 11,* 386–389.

Vernon-Feagans, L. (1997, April). Cumulative risk in understanding the developmental effects of otitis media: Attention, language, and social behavior. In I. Wallace (Chair), *Developmental sequelae of otitis media: New approaches to examining an old issue.* Symposium conducted at the Biennial Meeting of the Society for Research in Child Development, Washington, DC.

Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence.* New York: Psychological Corporation.

Weston, A. D., & Shriberg, L. D. (1992). Contextual and linguistic correlates of intelligibility in children with developmental phonological disorders. *Journal of Speech and Hearing Research, 35,* 1316–1332.

Winitz, H., & Darley, F. (1980). Speech production. In P. LaBenz & A. LaBenz (Eds.), *Early correlates of speech, language, and hearing* (pp. 232–265). Littleton, MA: PSG Publishing Co.

Zargi, M., & Boltezar, I. H. (1992). Effects of recurrent otitis media in infancy on auditory perception. *American Journal of Otolaryngology, 13,* 366–372.

Received March 3, 1998

Accepted July 22, 1999

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