CLINICAL LINGUISTICS & PHONETICS, 2003, VOL. 17, NO. 7, 529–547

A diagnostic marker for speech delay associated with otitis media with effusion: backing of obstruents

LAWRENCE D. SHRIBERG,
RAYMOND D. KENT, HEATHER B. KARLSSON,
JANE L. MCSWEENY, CONNIE J. NADLER
and ROGER L. BROWN

University of Wisconsin-Madison, WI, USA

(*Received 26 July 2002; accepted 19 September 2002*)

Abstract

A companion paper in this issue reported diagnostic accuracy findings for a marker (the *Intelligibility-Speech Gap*) to identify speech delay associated with otitis media with effusion (SD-OME). The present paper reports findings for another possible diagnostic marker for SD-OME—*Backing of Obstruents*. Conversational speech samples and citation forms from 48 speech-delayed children with positive and negative histories for OME were analysed using both perceptual and acoustic methods. The perceptual findings indicated significant trends for backing to be more prevalent in children with positive compared to negative histories of OME. Among a number of candidate speech error variables, lowered first spectral moments on lingual stops and sibilant fricatives as obtained from moments analysis emerged as a promising acoustic correlate of backing. Positive and negative predictive values, and sensitivity and specificity values for the acoustic marker, were above 75% for each of three stimulus words targeting /k/, /z/ and /ʃ/. Discussion considers alternative explanatory perspectives on the ontogenetic development of backing in children who have experienced the fluctuant hearing loss associated with early recurrent OME.

Keywords: Articulation, classification, hearing, phonology, speech disorder.

Introduction

A companion paper reviewed the central role of diagnostic markers in the assessment and explication of disease and disorder (Shriberg, Flipsen, Kwiatkowski

Address correspondence to: Lawrence D. Shriberg, Phonology Project, Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, WI 53705, USA. e-mail: shriberg@waisman.wisc.edu; website: http://www.waisman.wisc.edu/phonology/index.htm

Clinical Linguistics & Phonetics
ISSN 0269-9206 print/ISSN 1464-5076 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0269920031000138132

and McSweeny, 2003). The paper included diagnostic accuracy values for a behavioural marker termed the *Intelligibility-Speech Gap*, which was developed to identify a subtype of speech delay termed *Speech Delay-Otitis Media with Effusion* (*SD-OME*). The present paper reports diagnostic accuracy findings for a second proposed speech marker of SD-OME termed *Backing of Obstruents*. Operational definitions for Backing of Obstruents—articulating stops and fricatives at a more posterior lingual position—are provided in a later section. To gain a clear perspective on the findings to be reported, it is useful to begin with a brief historical overview of the conceptual origins of diagnostic markers for speech delay and to summarize relevant empirical findings for markers of SD-OME.

Preliminary studies

Over two decades ago a case study report described a pattern of atypical speech errors in a child whose speech delay was posited to be causally associated with the hearing loss that may occur during episodes of otitis media with effusion (Shriberg and Kwiatkowski, 1980; cf. 'Toby'). The report was one of several case studies contrasting articulation errors viewed as phonologically and phonetically natural (i.e. natural phonological processes) with those that were not natural. *Natural* sound changes are those attested in toddlers and preschool children acquiring the speech patterns of their language at normal rates (Ingram, 1976). The hypothesis proposed in the case study series was that sound changes that were *not* included in Ingram's synthesis of natural phonological processes might arise due to specific perceptual, cognitive, structural, speech-motor or affective deficits. Such proximal deficits, albeit more subtle than those observed in children with speech disorders of known origin (e.g. cleft palate, cerebral palsy, autism), might, in turn, provide clues to the distal causes or etiologies of speech delay of unknown origin. Thus, the concept of *non-natural* phonological processes was suggested as a rich source of potential diagnostic markers for etiological subtypes of speech delay.

The case history series was followed by empirical studies of an analysis procedure to assess both natural and non-natural sound changes in conversational speech samples. Findings based on samples from 110 children with typical and delayed speech acquisition indicated that a set of eight natural phonological processes accounted for approximately 92% of the children's sound changes (Shriberg and Kwiatkowski, 1983). Again, the naturalness of these eight classes of sound change (assimilation, cluster reduction, final consonant deletion, liquid simplification, palatal fronting, stopping, velar fronting and unstressed syllable deletion) could be motivated by appeal to their cognitive, perceptual, structural or speech-motor attributes. The approximately 8% of remaining sound changes across children were posited to reflect a constraint in one or more of these four biobehavioural domains—with identification of the specific domain (i.e. the proximal cause) providing important clues to the distal or etiological cause of the speech delay.

A follow-up report included speech data for 11 children with speech delay and significant histories of OME and 11 controls (speech delay and negative OME [OME−] histories), plus a cross-validation group of 15 children with significant OME and 40 controls (speech delay and OME− histories) assessed in a collaborative study with Dr. Barbara Hodson (Shriberg and Smith, 1983). For both data sets, the speech transcripts were coded for the same two classes of non-natural sound changes as reported in the 1980 case study. The first class, termed *Initial Consonant Sound Changes*, included three types of changes not typically observed in children with speech delay of unknown origin: deletion of initial consonants, replacement of initial consonants with /h/, and replacement of initial consonants by a glottal stop. The second class of purportedly non-natural sound changes, termed *Nasal Sound Changes*, also included three types of sound changes not typically observed in normal or delayed speech acquisition: replacement of a nasal by another nasal or stop, addition of an epenthetic stop before a nasal and denasalization. Statistically significant between-group differences were reported for both group comparisons, but a statistical limitation noted after publication (Dorothy Bishop, personal communication) tempers the interpretation of two of the inferential statistical findings.

Auditory-perceptual constraints were proposed to account for the six types of initial consonant and nasal sound changes described in 1980 and 1983. The thesis proposed to explain the initial sound changes, particularly deletion of prevocalic stops that were correct in intervocalic and post-vocalic positions, was that the intensity deficits in fluctuant conductive hearing loss would make it difficult to reliably discriminate the F2 locus for consonants in word-initial position (i.e. compared to cues for the place locus provided by the preceding vowel in intervocalic and post-vocalic consonants). Auditory-perceptual rationale for nasal sound changes—an early-acquired sound class typically well preserved in children with speech delay—was based on discrimination difficulties associated with the characteristic low-frequency energies in nasals (i.e. nasal murmur). Thus, the two sets of speech production errors were viewed as mirroring the perceptually mediated representation of speech sounds associated with fluctuant conductive hearing loss. Specifically, an average hearing loss of 25–30 dB in the speech frequencies is considered mild relative to average conversational levels. However, fluctuant loss at these levels (such as that which may occur during episodes of OME) was posited to be sufficient to perturb the earliest stages of a child's representation of the ambient phonology.

Cross-validation studies

In the remainder of the 1980s there were mixed findings from studies assessing the validity of these hypotheses on the two classes of non-natural sound changes. In an unpublished paper using methods similar to those used in the 1983 report, three children with speech delay and significant OME+ histories had more nasal distortions and more lateral sibilant distortions than 31 control children with speech delay and OME− histories (Kwiatkowski and Shriberg, 1983). However, in a retrospective study of 35 typically developing 3-year-old children (unselected for speech delay) who had been followed since birth at a paediatrics clinic, neither of the two classes of sound change were more frequent in the 16 children with significant OME+ histories compared to the 19 children with OME− histories (Shriberg, Kwiatkowski, Block, Katcher, Kertoy and Nellis, 1984). In a subsequent unpublished study, eight children with speech delay and significant OME+ histories made more of the non-natural sound changes than any of the 69 children with speech delay in two comparison groups (Shriberg and Kwiatkowski, 1985). During this time period, Naas and Loucks (1983) and Dyson, Holmes and Duffitt (1987) used perceptual measures and Kertoy (1983) used acoustic measures, to report descriptive support for the nasal sound changes in children with speech delay and significant OME+ histories. However, neither sound change category was statistically cross-validated in three larger scale studies of suspected SD-OME using diverse perceptual methods to sample, transcribe and analyse speech errors (Bishop and Edmundson, 1986; Paden, Novak and Beiter, 1987; Churchill, Hodson, Jones and Novak, 1988).

Recent trends

As reviewed in the companion paper, trends in the 1990s and into the present century have been toward two types of OME studies, neither of which are designed to identify possible diagnostic markers for the putative subtype of speech delay referenced here as SD-OME. A number of large scale epidemiologic studies have sought to assess whether the hearing loss associated with early frequent OME is a risk factor for short- or long-term deficits in speech-language acquisition or other areas of development. As summarized in a recent symposium, findings are not uniformly negative but are skewed toward the general conclusion that OME is not a risk factor for significant developmental delay (Roberts and Hunter, 2002). On balance, however, the other class of studies, which focuses on detailed descriptions of children's babbling and early word forms, does provide substantial support for the influence of hearing loss on the acquisition of speech (cf. Roberts and Hunter).

Statement of the problem

The study to be reported addresses a third class of questions in the OME-speech literature, one that resumes the line of research on potential diagnostic markers for SD-OME pursued in the reviewed studies ending in the late 1980s. The goal of identifying diagnostic markers for etiological subtypes of speech delay seems particularly appropriate at a time when substantive and technical advances make it possible to pursue questions about the origin and nature of disease and disorder. The goal of the present study was to determine if there was a non-natural speech change that met conventional clinical criteria for a diagnostic marker of speech delay suspected to be consequent to early recurrent otitis media with effusion.

Method

Participants

Table 1 is a summary of descriptive information for a sample of 48 children with speech delay of currently unknown origin. Speech samples and case records of all children were selected from three prior studies of child speech-sound disorders reported in Shriberg and Kwiatkowski (1994) and in Shriberg, Gruber and Kwiatkowski (1994). The screening criterion for possible inclusion in these studies was the presence of age-level discrepancies in speech production as sampled in conversational speech (Shriberg, Austin, Lewis, McSweeny and Wilson, 1997). Conventional exclusionary criteria included evidence supporting any sensory, cognitive, structural, speech-motor or affective impairment that would warrant an alternative diagnostic label.

Four additional requirements for inclusion in the present study focused on a

Table 1. *Description of the 48 participants divided into four groups based on their histories of early recurrent otitis media with effusion (OME) and their scores on the Percentage of Consonants Correct (PCC) and Intelligibility Index (II) (Shriberg* et al.*, 1997)*

Group				Age (months)			Sex		PCC (%)			II (%)		
Number	*n*	OME history	II status	*M*	*SD*	Range	% Male	% Female	*M*	*SD*	Range	*M*	*SD*	Range
1	12	positive	low	43.2	5.1	36–53	66.7	33.3	59.5	10.3	38.8−72.7	73.1	12.8	41.5−86.4
2	12	positive	high	56.1	6.7	46–67	66.7	33.3	66.9	8.1	50.6−78.1	96.6	1.5	93.5−99.1
3	12	negative	low	45.0	6.4	35–56	58.3	41.7	60.3	10.3	38.7−71.5	78.9	6.4	67.6−86.4
4	12	negative	high	53.0	6.1	46–66	58.3	41.7	66.9	8.1	52.5−77.0	96.7	2.1	93.6−99.5
Total	48			49.3	8.0	35–67	62.5	37.5	63.4	9.7	38.7−78.1	86.3	12.7	41.5−99.5

child's history of early recurrent otitis media with effusion, intelligibility scores, age and gender. The primary criterion for inclusion in a subgroup termed OME+ was a history indicating that the child had experienced at least five episodes of OME, with strong parental support data indicating the likelihood of hearing loss during some or all of these episodes. A total of 24 children were identified who reportedly 'didn't hear as well' during episodes of OME, with hearing loss for many children documented by audiological examination reports. To maximize the clinical generalizability of the present findings, the emphasis for the present study was on this type of case history data, typically obtained from parents during intake interviews and follow-up records requests to other health care professionals. Thus, unreliability in such data would increase both false positives and false negatives, and decrease the likelihood of rejecting the null hypothesis of no between-group differences.

The second inclusionary criterion for the current study was that participants have either relatively high or relatively low intelligibility in conversational speech. The decision to stratify participants by intelligibility status was based on findings in the companion paper indicating that intelligibility status may be a significant correlate of SD-OME. The intelligibility metric used for this purpose was the percentage of words that a transcriber could confidently gloss from an audio recording of a child's conversational speech with a trained examiner (Shriberg *et al.*, 1997). Transcribers were permitted three or more repetitions of a string of speech to gloss what the child intended to say and were aided by the gloss and comments on linguistic context recorded on-line by the examiner. The transcriber was permitted to agree or disagree with the examiner's gloss, using a set of asterisk conventions to indicate unintelligible syllables and words (Shriberg, Allen, McSweeny and Wilson, 2001). For the present study, Intelligibility Index percentages greater than 93% were classified as High Intelligibility (HI) and those lower than 87% as Low Intelligibility (LI). These cut-off criteria were used to maximize intelligibility differences in the available sample pool. Reference data on the Intelligibility Index indicate that children with speech delay in the age range of this study (approximately 3–6 years) average 86.9% (SD = 11.7%) intelligible words in conversational speech (Austin and Shriberg, 1996). The third inclusionary criterion was a sampling pass that attempted to optimally balance groups by age and gender.

As shown in table 1, the 48 participants were comprised of four 12-participant groups: Group 1: OME+ and Low Intelligibility, Group 2: OME+ and High Intelligibility, Group 3: OME− and Low Intelligibility and Group 4: OME− and High Intelligibility. One-way analyses of variance and a chi-square test indicated that there were no significant ($p < 0.05$) between-group differences in age, sex or Percentage of Consonants Correct (PCC) scores. As above, participants ranged in age from 2 years, 11 months to 5 years, 7 months. The overall sex proportion (30 males, 18 females = 63% males) was consistent with sex ratios reported for preschool-age children with speech-sound disorders (Shriberg and Kwiatkowski, 1994). Intelligibility Index scores, used to differentiate groups 1 and 3 from groups 2 and 4, ranged from approximately 42% to 100%. Percentage of Consonants Correct scores ranged from approximately 39% to 78%, indicating that these children had moderate-severe to severe speech delay (Shriberg *et al.*, 1997). Based on a series of independent estimates, these two most involved severity levels account for 42% of children with speech delay referred to a university speech clinic (Shriberg and Kwiatkowski, 1994).

Assessment

The recorded samples used for the present study were obtained from two speech tasks. All recordings had been obtained using a Sony 5000 monaural audiocassette recorder and a matching remote microphone monitored at a lip-to-microphone distance of approximately 15 cm. Recording procedures included well-developed conventions to maximize signal-to-noise ratios.

The first speech task was a conversational speech sample in which children were invited to converse about topics such as their daily activities, friends, and past and upcoming special occasions. The second task was a modified administration of the *Photo Articulation Test* (*PAT*: Pendergast, Dickey, Selmar and Soder, 1984). The research protocol used in the studies from which these participants were selected directed the examiner to obtain two responses to each PAT stimulus word. The examiner first attempted to evoke a response spontaneously, using the pictures in the test book. If spontaneous evocation was not successful, the examiner used delayed imitation (e.g. 'That's a __; say the word') to evoke the correct word. The examiner also obtained a second response to each word using direct imitation. Thus, the data set available for the present analysis included spontaneous responses for most words and imitative responses for all words.

Perceptual analysis

Preliminary analyses of both the perceptual and acoustic data indicated that, with the exception of one sound change, the findings associating positive OME history with atypical sound changes failed to reach customary levels of statistical significance. The one sound change that did appear to warrant thorough perceptual and acoustic analysis was backing of lingual stops and sibilant fricatives. This sound change had been added as a third non-natural sound change class to the two non-natural sound change classes described in the previous review, based on three observations: more lateralized (i.e. perceptually similar to backing) sibilants in children with suspected SD-OME (Kwiatkowski and Shriberg, 1983), a colleague's informative observation of backing in a child with a history of significant OME (Dr. Barbara Hodson, personal communication), and the first author and a colleague's (Joan Kwiatkowski's) frequent observations of backing in children seen clinically who had significant OME histories.

A software utility (Shriberg, Lof and Wilson, 1993) was used to identify and statistically compare the frequency of occurrences of backing in the 24 children with positive histories of OME to the frequency of backing in the 24 children with negative histories of OME. This analysis was obtained for each of the three sets of data available for the 48 participants: the conversational speech sample, the spontaneous (or delayed imitation) PAT responses and the imitative PAT responses. Additional procedural information on these analyses is deferred to the discussion of findings in Results.

Acoustic analyses

Preliminary analyses

Acoustic analyses of backing in the two groups of speakers were completed on a set of 21 PAT words, for which spontaneous (i.e. citation) responses had been obtained. The analyses were preceded by two scans through the raw data to eliminate responses that would not be appropriate for further analysis. First, tokens were eliminated if the wrong word had been produced, or if examiner overtalk or extraneous environmental noise had obscured the target sound(s) in the word. Data loss from such factors was approximately 20% (i.e. 196 of the original 1008 [48×21] tokens were eliminated).

The second exclusionary analysis was based on a review of the narrow phonetic transcription of each participant's responses. Responses were classified into seven categories based on the participant's production of the target segment in the word as symbolized in narrow phonetic transcription. The seven categories identified whether the target was (1) produced correctly, (2) replaced by another consonant of the same manner, but at a more posterior place (i.e. backed), (3) backed, as indicated by a backing diacritic, (4) fronted, as indicated by a fronting diacritic, (5) deleted, (6) replaced by a consonant of a different manner or (7) replaced by a vowel. Only those responses meeting criteria for any of the first four categories were retained for analyses. The result of this analysis was the elimination of an additional 140 responses of the 812 remaining eligible responses. Thus, the data set appropriate for acoustic analyses included 672 or 66.7% of the original 1008 responses to the 21 PAT words.

Table 2. *Citation forms and acoustic measurements included in the analyses of stop and fricative productions of the 48 participants*

				Moments	
Class	Target Feature	Sound	Word	Burst	Frication
Stop					
	Alveolar	/t/	table	X	
			teeth	X	
			TV	X	
	Velar	/k/	book	X	
			cake (2)	XX	
			can	X	
			car	X	
			cat	X	
			comb	X	
			cup	X	
		/g/	egg	X	
			flag	X	
Fricatives					
	Alveolar	/s/	sandwich		X
			saw		X
			scissors		X
			house		X
		/z/	zipper		X
			keys		X
			scissors		X
	Palatal	/ʃ/	shoe		X
			brush		X
			fish		X
Totals			21	13	10

Table 2 is a list of the 21 PAT words that yielded data on six target sounds for the acoustic analyses. Two words (*cake, scissors*) yielded sufficient numbers of tokens for analyses of two consonant sounds. The target sounds included one alveolar stop (/t/), two velar stops (/k/, /g/), two alveolar fricatives (/s/, /z/) and one palatal fricative (/ʃ/). Each of these six obstruents can be articulated with more posterior lingual positioning, as indexed acoustically by the first spectral moment within the stop burst or at the temporal middle of the fricative noise. As shown in the totals at the bottom of table 2, there were slightly more stop (13) than fricative (10) targets.

Procedures

Acoustic analyses of the stop and fricative segments of each of the words in table 2 (as produced by the 48 participants) were completed. Eligible responses were digitized using a Tascam 112MK II tape deck connected to Computer Speech Lab 4300B (CSL4300B) hardware. The signal was sampled at 20 kHz with 16 bits of quantization using the CSL 4300B record facility.

Acoustic analyses included wide-band (400 Hz) spectrograms with superimposed LPC formant history (CSL 4300B) and spectral moments computed from FFT spectra (CSpeech). The CSL analyses were performed with a 10 ms frame length, 10 ms frame advance, and 16-coefficient filter order for LPC. Stops were identified by the presence of a burst and aperiodic aspiration noise in the region of 4000–6000 Hz. In order to centre the first 20 ms Hamming window on the burst, the left cursor was placed 10 ms before the onset of the stop burst; the right cursor was placed at the end of aspiration noise. Fricatives were identified by the presence of strong aperiodic energy as evidenced in both the waveform and the spectrogram. The left cursor was placed at the onset of strong aperiodic energy, and the right cursor was placed at the offset of strong aperiodic energy. Moments files were subsequently created and analysed in CSpeech with a 20 ms Hamming window and 10 ms step.

The moments files were placed into spreadsheets for further analysis. The middle time slice (i.e. the middle 20 ms) was identified and isolated for the fricatives. In cases where a file contained an even number of analysis windows, one of the two middle windows was selected randomly. The initial time slice (i.e. the burst) was identified and isolated for the stops.

Reliability

The papers from which the present samples were obtained include information on the procedures used to assess transcription reliability for the transcribers who completed consensus transcriptions on the conversational speech samples and the PAT words. Point-to-point agreement for narrow phonetic transcription ranged from 79% to 82% (Shriberg and Kwiatkowski, 1994), which is consistent with values reported for transcription in child phonology studies.

An estimate of the reliability of the acoustic data was obtained by resegmenting and measuring a randomly drawn sample of 77 stops and fricatives (12%) from the 672 eligible target words. Over all tokens, the average difference between the two measures of mean frequency was 52 Hz (SD = 164 Hz). For stops ($n = 53$) the mean difference was 25 Hz (SD = 39 Hz), and for fricatives ($n = 24$) the mean difference was 110 Hz (SD = 283 Hz). These agreement values are consistent with values reported in the acoustic literature and are considered adequate for the effects to be reported.

Results and discussion

Perceptual findings

Analyses

Analyses of both the conversational speech samples and the spontaneous and imitative responses to the PAT stimuli shown in table 2 were completed using routines in the PEPPER suite. This software provides information on error patterns subtotalled for several structural-linguistic variables. Figure 1 provides a graphic summary of the perceptual data. Panel A includes information on the percentage of children who backed stops and/or fricatives in conversational speech, Panel B provides these data as produced spontaneously in naming the 21 PAT words, and Panel C provides the backing data for PAT words evoked by imitation. Within each panel, comparative data for the stops, fricatives, and the combined stop and/or fricative targets are provided for the 24 participants in the OME+ groups (dark grey) and the 24 participants in the OME− groups (light grey). Information under each comparison indicates each sound target's position in the PAT word (Initial: *I*, Medial: *M*, Final: *F*, All positions: *A*), structural context (Singleton: *S*, Cluster: *C*, Both singleton and cluster: *B*), and the number of occurrences of the sound change in the sample (1–4+ occurrences). For clarity, figure 1 includes only those OME+ versus OME− differences that were statistically significant ($p < 0.05$). One-tailed significance levels were used, because the prediction was that more participants in the OME+ than OME− group would back stops and/or fricatives. Thus, the comparisons in figure 1 provide a profile of the findings for backing at the level of individual participants.

Findings

The pattern of findings in figure 1 indicates that the percept of backed lingual stops and sibilant fricatives occurred in a higher percentage of speech delayed children with positive compared to negative histories of OME. There were 84 such comparisons tested (28 comparisons $\times$ three sample types). As indicated in figure 1, 27 (32%) of the individual comparisons (i.e. excluding the 'stops and/or fricatives' combined comparisons) were statistically significant at the 0.05 alpha level or higher. As shown in figure 1, all 28 comparisons indicated that a higher percentage of children in the OME+ than OME− group had from 1 to 4 or more occurrences of backing as obtained in the three types of speech samples. Some additional trends in figure 1 warrant discussion for methodological and conceptual issues in this and future research.

From a methodological perspective, the data in figure 1 indicate that citation forms may be a more sensitive sampling environment for diagnostic markers of SD-OME, and by extension, possibly for diagnostic markers of other subtypes of speech delay. Excluding the set of comparisons in which a child may have backed 'stops and/or fricatives', there were over twice the number of significant comparisons in the spontaneous (19) and imitative (17) PAT word samples compared to those obtained in the conversational speech samples (eight). These

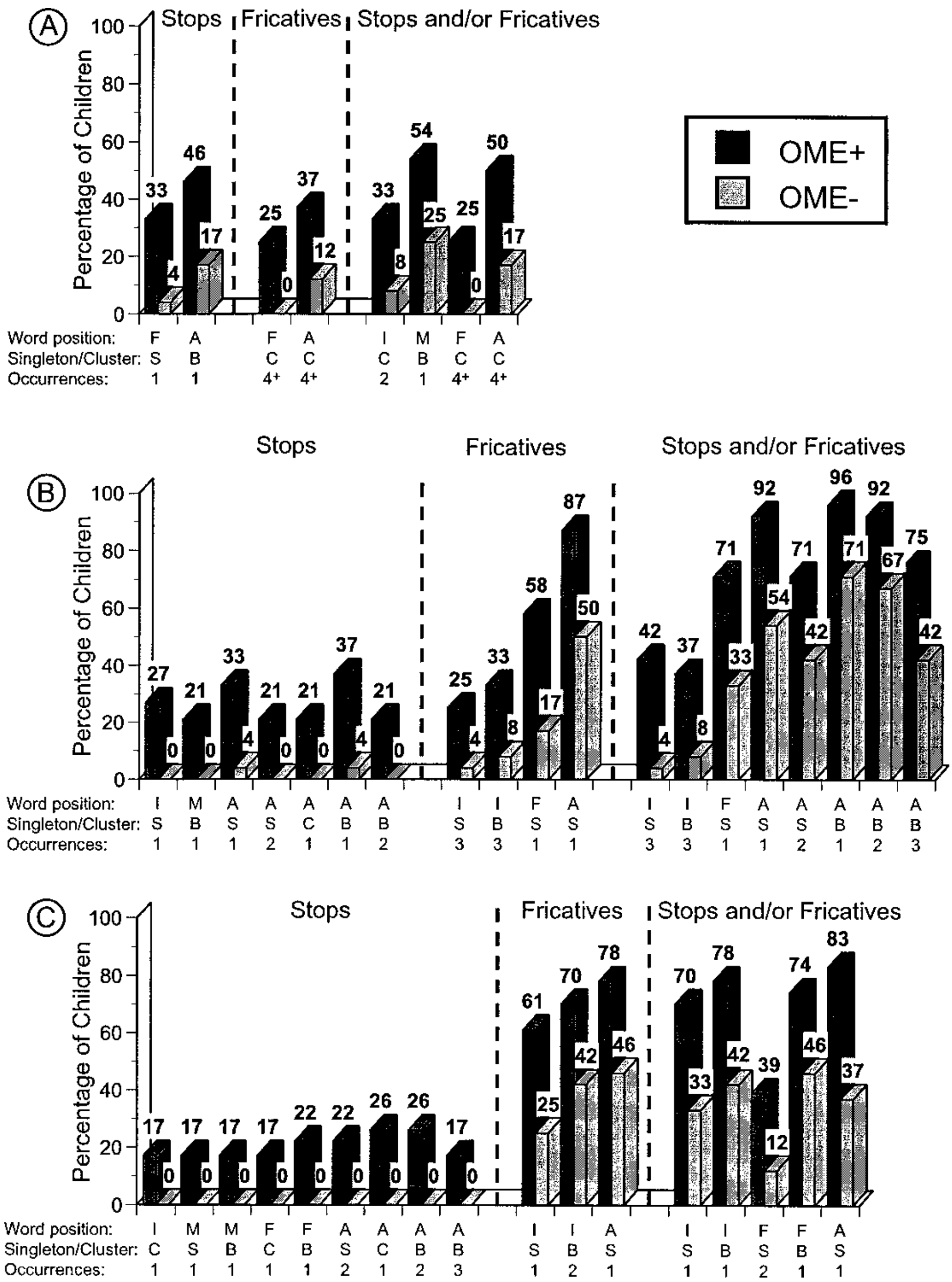

Figure 1. *Summary of the percentage of children who backed stop and/or fricative obstruents in conversational speech (Panel A), in spontaneous responses to the Photo Articulation Test (PAT) (Panel B), or in imitative responses to the PAT. See text for a key to the abbreviations on the abscissa.*

differences could be due to one or more of the following factors. First, more children may, in fact, back in the context of citation forms versus in conversational contexts. Second, transcribers may be more sensitive to backing in citation forms versus in conversational speech. Third, the generalizability of these findings could be constrained by some procedural or linguistic characteristics of the citation forms and conversational samples used in the present study. If the first interpretation is cross-validated as correct, such findings suggest that citation form testing may be more useful than conversational sampling for the types of allophone-level information needed in some diagnostic markers proposed for child speech-sound disorders.

From a descriptive linguistic perspective, the findings in figure 1 also suggest that backing is widely distributed relative to the word position and the structural contexts in which it occurs. Although tallies indicate that significant between-group comparisons occurred more often for singleton targets in word-initial (prevocalic) position, backing also occurred in consonant clusters and for target sounds in word-medial (intervocalic) and word-final (postvocalic) positions.

The data in figure 1 also indicate that mode of evocation for citation forms was not associated with the frequency of significant comparisons in the occurrence of backing in participants with OME+ versus OME− histories. Specifically, spontaneous naming of the 21 PAT words yielded 19 significant comparisons (Panel B) compared to 17 for the imitative responses (Panel C). Moreover, the magnitudes of significant between-group comparisons were also essentially similar when the PAT words were evoked spontaneously rather than by imitation. One possible explanation for the similarity in the occurrences of backing in different sampling modes could be a perceptual bias in transcription. The two responses obtained to the same PAT word were not randomly transcribed, but rather followed one another on the tape. It should be noted, however, that high stability across sampling modes is typical of distortion-type errors as sampled in children with speech delay. That is, the provision of imitative cues to a child with distortion errors is not typically sufficient for the child to readily increase articulatory precision. However, provided that a target sound is in a speaker's phonetic inventory, he or she may profit from the cues provided by an imitative stimulus and not delete the target or substitute another sound for it.

Finally, the data in figure 1 indicate that as defined in this study, backing occurred relatively frequently in children with positive histories of OME. An average of the percentage of participants in each OME status group with 1 to 4+ occurrences of backing across the 44 significant comparisons yielded a 28% difference. Tallies indicated that approximately 45% of participants in the OME+ group backed a stop or fricative at least once, compared to an average of approximately 17% of the participants who backed obstruents at least once in the OME− group.

The findings have implications for theory and method in otitis-speech research. An eventual account of the origins of backing will need to include explanation for its occurrence (albeit in fewer cases) in children whose speech delay may be associated with other perceptual, cognitive, structural or speech-motor constraints. Two possible methodological explanations for its occurrence in children with OME− histories are that these histories were misclassified (i.e. false negatives), or that children's speech production was inaccurately transcribed. The present design does not allow for a test of the possibility of misclassification, because controlled audiological data for participants in both OME status groups were not available. However, the findings to follow from the acoustic analyses do provide an assessment of the reliability of findings based on narrow phonetic transcription. If the present findings are cross-validated, they might be interpreted as suggesting that only approximately 50% of children with OME+ histories have had the frequency and degree of hearing loss that leads to the backing of obstruent consonants.

Acoustic analyses

Preliminary analyses

The perceptual analyses indicated that children with OME+ histories backed obstruents more often than children with OME− histories. The goal of a series of acoustic analyses was to assess the ability of an acoustic marker of backing to accurately discriminate children's OME histories (i.e. to 'diagnose' the presence of an OME+ history). For these purposes, the acoustic analyses were limited to words said by all participants, a criterion that was met by using the spontaneous responses to the 21 PAT words listed in table 2. A series of preliminary analyses was completed to determine which linguistic and subject factors might be collapsed to provide more statistical power. Preliminary descriptive analysis of the frequency data indicated that the analyses should not be collapsed across place-manner groups, but rather should be completed individually for the four feature classes shown in table 2.

To determine whether it was appropriate to treat each word separately, four three-way analyses of variance of the spectral data were completed. The four analyses included the words listed in table 2, divided into the four manner-place categories: the alveolar stop /t/ (three words), the two velar stops /k/ and /g/ (nine words; ten tokens), the two alveolar fricatives /s/ and /z/ (seven words) and the palatal fricative /ʃ/ (three words). Each statistical model included main effects for group (OME+, OME−), intelligibility (HI, LI), stimulus words (3–10), and all two-way and three-way interactions among the three main factors.

The pattern of findings from the four analyses yielded no significant main effects for OME or intelligibility, but did yield significant two-way and three-way interactions for each of the main factors with stimulus words. For the present purposes, the diagnostic marker analysis was therefore completed at the word level. A problem with this approach is that the limited number of participants in each cell of the design (12) and the token loss per word (see Method) provided relatively few data points (ranging from 7–21 for the groups divided by intelligibility status) on which to base each analysis. Confidence intervals, which are heavily weighted by cell sizes, were expected to be large for these data.

Acoustic findings

Figure 2 is a display of the first spectral moment (M1) findings for the obstruent sounds in the 21 words, arranged by each of the two intelligibility levels and by sound class. Panels A and B include, respectively, M1 values for the fricatives produced by the 12 children with high intelligibility and the 12 children with low intelligibility. Panels C and D include, respectively, M1 values for the stops produced by the 12 children with high intelligibility and the 12 children with low intelligibility. To aid interpretation, the words within each place-manner class in each panel are arranged in increasing order relative to the average frequency of the target stop burst or fricative noise. The referent group for this sort is the data

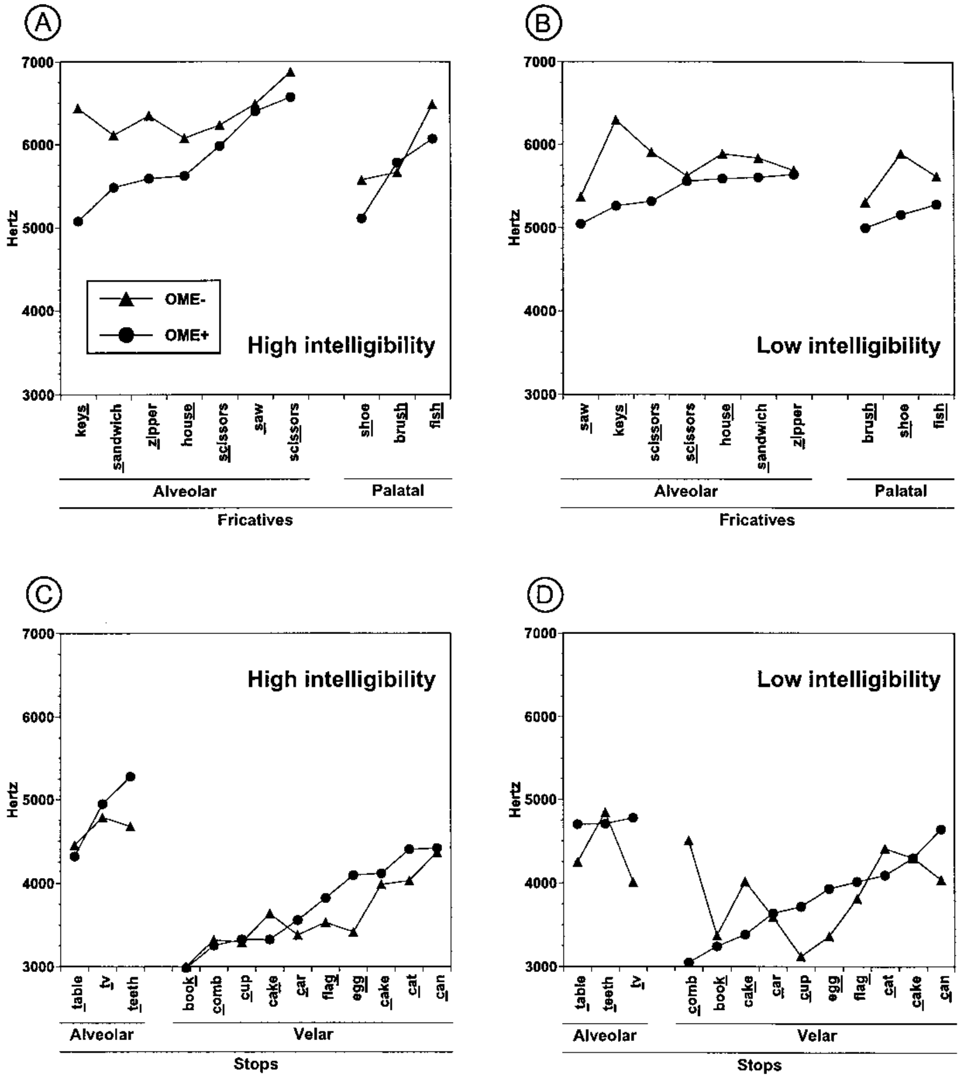

Figure 2. *First spectral moment (M1) findings for the obstruent sounds in the 21 words, arranged by each of the two intelligibility levels and by sound class. Panels A and B include, respectively, M1 values for the fricatives produced by the 12 children with high intelligibility and the 12 children with low intelligibility. Panels C and D include, respectively, M1 values for the stops produced by the 12 children with high intelligibility and the 12 children with low intelligibility.*

points from the OME+ group (circles); in comparison to the OME− group (triangles), the OME+ group was expected to have lower M1 values on obstruents.

Although the previous statistical analyses indicated that the M1 values of the OME+ versus OME− groups across all target words were not significantly different, the data in the top panel in figure 2 suggest a clear trend for children with OME+ histories to have lower M1 values on the target fricatives. As indicated in panels A and B, these directional trends were obtained for 19 of the 20 comparisons (i.e. there was a reversal in direction for /ʃ/ in the word *brush*).

A second observation on the data in panels A and B is the trend for children in the low intelligibility group to have lower M1 values for fricatives compared to the values for children in the high intelligibility group. Again, the analysis of variance did not indicate that this difference was statistically significant, which, at least in part, could be associated with limitations in statistical power. It is interesting to speculate about possible relationships between fricative backing and lowered intelligibility. Behaviours in the two domains may be associated in some correlative or possibly causal fashion, either of which is of interest for explanatory models relating fluctuant hearing loss to intelligibility and to the acquisition of precise allophones. One possibility is that backed tongue postures interfere with the optimum acoustic features for intelligible speech in American English. That is, backing may be associated with other articulatory-acoustic correlates not addressed in this study, and these features may reduce the clarity of vowels and thereby reduce intelligibility.

Finally, the data for the alveolar and velar stops in panels C and D are not consistent with the perceptual data for stops as displayed in figure 1. That is, the interleaving trends in the acoustic data do not mirror the clear trends in the perceptual data for participants with OME+ histories to frequently back stops. It is not clear whether the differences between the perceptual and the acoustic data are valid, or whether they might reflect a methodological problem with the acoustic parameter selected to index backing. On the latter possibility, acoustic analyses of alternative potential correlates of backing of stops, and individual comparisons for children across perceptual and acoustic measures, are beyond the scope of the present study. A difficulty in an acoustic study of backing is that the acoustic correlates of this feature may vary to some degree with the place and manner of the target consonants. Therefore, no single acoustic property may be uniformly sensitive to backing across consonants that differ in place and manner. This problem may be particularly serious with the velar stop, which has a variable place of articulation determined principally by vowel context. It may be necessary to study a larger number of tokens in each place-manner category to identify the most useful acoustic correlates of backing. It appears from figure 2 that the first spectral moment may better distinguish OME+ and OME− for some phonetic contexts than others. Note that the largest differences occur in panels A and B for alveolar fricatives that occur in a front-vowel environment (e.g. *keys, sandwich, scissors, zipper*).

Diagnostic marker analysis

A third analysis was completed to determine if responses to any of the 21 words met criteria for diagnostic accuracy in discriminating a participant's OME history. Table 3 is a summary of findings for the three words that met a minimum criterion of 75% accuracy on each of five conventional clinical diagnostic metrics: positive predictive value, negative predictive value, sensitivity, specificity and diagnostic accuracy (averaged sensitivity and specificity). Findings for *shoe* were included, because sensitivity was close to criteria (0.71) with acceptable to high values on the other four metrics. Included in table 3 are the mean M1 values for OME+ and OME− participants in each comparison, and accuracy values (with confidence intervals) based on logistic regressions of the M1 values for each of the two OME history groups. Because of the expected large confidence intervals for the prediction

Table 3. *Summary of the acoustic findings for backing of obstruents as a diagnostic marker for speech delay associated with otitis media with effusion*

					Mean M1[a] Values (Hz)				Diagnostic Metrics and Confidence Intervals[b]				
Stimulus Word	Target Sound	Place-Manner	Word Position	Intelligibility Status	OME−	SE[c]	OME+	SE[c]	Positive Predictive Value	Negative Predictive Value	Sensitivity	Specificity	Diagnostic Accuracy
can	/k/	Velar Stop	Initial	Low	4033	243	4638	272	0.75 (0.45–1.1)	0.80 (0.55–1.1)	0.75 (0.45–1.1)	0.80 (0.55–1.1)	0.78 (0.59–0.97)
keys	/z/	Alveolar Fricative	Final	Low	6297	348	5261	426	0.75 (0.33–1.2)	0.83 (0.54–1.1)	0.75 (0.33–1.2)	0.83 (0.54–1.1)	0.80 (0.55–1.1)
shoe	/ʃ/	Palatal Fricative	Initial	Low	5892	178	5160	178	0.75 (0.45–1.1)	0.83 (0.54–1.1)	0.86 (0.60–1.1)	0.71 (0.38–1.1)	0.79 (0.51–1.0)

[a]M1: First spectral moment.
[b]Confidence intervals for the classification rule for each logistic regression are shown in parentheses.
[c]SE: Standard error of the mean.

rule generated by the logistic regression, which are in part associated with the small cell sizes in this study, the discriminant rules were not included in the table.

As shown in table 3, differences in the average frequency of the stop burst in /k/ were sufficiently large to discriminate the OME+ histories of speakers with low intelligibility as defined for this study. However, the directional difference between means was opposite to that predicted for backing; as above, additional research is needed to identify the appropriate correlate of perceptually backed stop bursts.

For the remaining two sounds in table 3, the diagnostic accuracy data indicated that the production of /z/ in *keys* and /ʃ/ in *shoe* met preliminary criteria as a diagnostic marker of children's OME histories. As shown in figure 2, /z/ in *keys* was associated with maximum separation of the averaged mean frequencies for speakers in the two OME history groups (over 1,000 Hz). It is interesting to note that the vowels in these three words include three of the four point-vowels on the vowel quadrilateral—the low front vowel /æ/ in *can*, the high front vowel /i/ in *keys*, and the high back vowel /u/ in *shoe*. These findings would suggest that backing is not closely tied to common coarticulatory processes such that children who tend to back consonants will more likely do so in specific phonetic contexts. On this point, notice that there does not seem to be a discernable pattern to the ordering of words in figure 2 relative to vowel place and height.

Conclusions

The findings of this study are viewed as support for the possibility of identifying diagnostic markers for children with speech delay of currently unknown origin. The etiological subtype of speech delay of interest in the present study includes those children whose speech delay may be consequent to the fluctuant hearing loss associated with early recurrent otitis media with effusion. The hypothesis predicts that the degree and type of conductive hearing loss associated with OME+ can affect a child's ability to establish the correct allophones for some stop and fricative consonants. Specifically, the fluctuant hearing loss makes it difficult to reliably perceive and thus establish correct place for obstruents, yielding a more posterior tongue placement for at least sibilant fricatives. Support for this general hypothesis is based on both the perceptual and acoustic data reviewed in this report, which included speech production sampled in three ways.

If cross-validated in continuing studies, implications for the present findings extend to both theory and practice. Theoretical issues underscore the impact of environmental variables on speech acquisition, with the possibility that slight articulatory differences can be traced to early difficulties in veridical perception of the speech forms of the ambient community. As discussed in Roberts, Burchinal, Koch, Footo and Henderson (1988), the possible effects of such seemingly minor perturbations early in the development of a system are consistent with the principles of self-organizing systems and dynamic systems frameworks. From this perspective, fine-grained longitudinal tracking of the time course of allophone-level sound changes (such as backed consonants) could provide a useful research paradigm for the study of child phonology. As a complement to the emerging interests and findings on the genetic contributions to typical and atypical speech-language acquisition (cf. Leonard, 1998; Stromswold, 1998; 2001), such non-genetic causal variables mediating perception should be of continuing research interest.

From a clinical perspective, because backing may not be perceptually evident to either the speaker or the untrained listener, it is possible that such slight differences could persist over a speaker's life span (cf. Lewis and Shriberg, 1994; Austin and Shriberg, 1996). Although not itself a behaviour with consequences for social or vocational function, the persistence of backing in children with speech delay may suggest the value of perceptually based treatment to highlight the salient acoustic cues for articulate speech. That is, specific focus on relevant perceptual cues would appear to be warranted for children whose allophonic-level differences implicate auditory-perceptual deficits, rather than cognitive, structural, speech-motor or affective-motivational issues as original and/or maintaining causal variables. The effectiveness of such perceptual approaches for children with suspected SD-OME would seem to be especially important to explore at a time when, as noted in the introduction to this issue, there has been a notable emphasis on oral motor rather than auditory-perceptual techniques for children with speech delay of unknown origin.

Acknowledgements

Our thanks to Gregory Lof and David Wilson for their important contributions to the software utility developed for the perceptual data in this study. Thanks also to the following colleagues for their assistance and counsel at different stages of this project: Peter Flipsen Jr., Katherina Hauner, Barbara Hodson, Marilyn Kertoy, Joan Kwiatkowski, Adele Miccio, Elaine Paden, Alison Scheer, Helen Thielke and Marie Wirka.

References

Austin, D. and 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).

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

Churchill, J. D., Hodson, B. W., Jones, B. W. and 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.

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

Ingram, D., 1976, *Phonological Disability in Children: Studies in language disability and remediation*, second edition (New York: Elsevier Publishing Company, Inc.).

Kertoy, M. K., 1983, Production of voiced stops and nasals in children with a history of middle ear involvement: an acoustic analysis. Unpublished manuscript, University of Wisconsin-Madison.

Kwiatkowski, J. and Shriberg, L. D., 1983, *Classification studies of developmental phonologic disorders: Evidence for subgroups*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Cincinnati, OH, November.

Leonard, L. B., 1998, *Children with Specific Language Impairment* (Cambridge, MA: MIT Press).

Lewis, B. A. and Shriberg, L. D., 1994, *Life span interrelationships among speech, prosody-voice, and nontraditional phonological measures*. Mini-seminar presented at the Annual Convention of the American Speech-Language-Hearing Association, New Orleans, LA, November.

NAAS, J. F. and LOUCKS, J. D., 1983, *Phonological change following otologic intervention*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Cincinnati, OH, February.

PADEN, E. P., NOVAK, M. A. and BEITER, A. L., 1987, Predictors of phonologic inadequacy in young children prone to otitis media. *Journal of Speech and Hearing Disorders*, **52**, 232–242.

PENDERGAST, K., DICKEY, S., SELMAR, J. and SODER, A., 1984, *The Photo Articulation Test*, revised edition (Dansville, IL: Interstate Printers and Publishers).

ROBERTS, J. E., BURCHINAL, M. R., KOCH, M. A., FOOTO, M. M. and 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. and HUNTER, L. (Chairs), 2002, *Otitis media and language learning sequelae: current research and controversies*. Symposium conducted in Arlington, VA, May.

SHRIBERG, L. D., ALLEN, C. T., MCSWEENY, J. L. and WILSON, D. L., 2001, PEPPER: Programs to examine phonetic and phonologic evaluation records [Computer software] (Madison, WI: Waisman Center Research Computing Facility, University of Wisconsin-Madison).

SHRIBERG, L. D., AUSTIN, D., LEWIS, B. A., MCSWEENY, J. L. and WILSON, D. L., 1997, The Percentage of Consonants Correct (PCC) metric: extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, **40**, 708–722.

SHRIBERG, L. D., FLIPSEN, P., JR., KWIATKOWSKI, J. and MCSWEENY, J. L., 2003, A diagnostic marker for speech delay associated with otitis media with effusion: the intelligibility-speech gap. *Clinical Linguistics and Phonetics*, **17**, 507–528.

SHRIBERG, L. D., GRUBER, F. A. and KWIATKOWSKI, J., 1994, Developmental phonological disorders III: Long-term speech-sound normalization. *Journal of Speech and Hearing Research*, **37**, 1151–1177.

SHRIBERG, L. D. and KWIATKOWSKI, J., 1980, *Natural Process Analysis: A procedure for phonological analysis of continuous speech samples* (New York: Macmillan).

SHRIBERG, L. D. and KWIATKOWSKI, J., 1983, *Classification studies of developmental phonologic disorders: Some methodological data*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Cincinnati, OH, November.

SHRIBERG, L. D. and KWIATKOWSKI, J., 1985, *Studies in early recurrent otitis media and speech delay*. Paper presented at the Annual Convention of the Wisconsin Speech-Language-Hearing Association, Madison, WI, April.

SHRIBERG, L. D. and KWIATKOWSKI, J., 1994, Developmental phonological disorders I: a clinical profile. *Journal of Speech and Hearing Research*, **37**, 1100–1126.

SHRIBERG, L. D., KWIATKOWSKI, J., BLOCK, M. G., KATCHER, M. L., KERTOY, M. K. and NELLIS, R. A., 1984, *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, November.

SHRIBERG, L. D., LOF, G. L. and WILSON, D. L., 1993, A utility program to aggregate non-natural sound changes [Computer software] (Madison, WI: Waisman Center Research Computing Facility, University of Wisconsin-Madison).

SHRIBERG, L. D. and 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.

STROMSWOLD, K., 1998, Genetics of spoken language disorders. *Human Biology*, **70**, 297–324.

STROMSWOLD, K., 2001, The heritability of language: a review and metaanalysis of twin, adoption and linkage studies. *Language*, **77**, 647–723.