Developmental Apraxia of Speech: II. Toward a Diagnostic Marker

Lawrence D. Shriberg
University of Wisconsin-Madison

Dorothy M. Aram
Emerson College
Boston, MA

Joan Kwiatkowski
University of Wisconsin-Madison

This second paper in a series on developmental apraxia of speech (DAS) (Shriberg, Aram, & Kwiatkowski, 1997a) reports findings from two studies. Study I compares speech and prosody-voice profiles of a group of 14 children with suspected DAS to profiles of 73 children with speech delay (SD). Results suggest that the only linguistic domain that differentiates some children with suspected DAS from those with SD is inappropriate stress. Study II cross-validates these findings, using retrospective data from a sample of 20 children with suspected DAS evaluated in a university phonology clinic over a 10-year period. Discussion considers methodological and conceptual issues in the measurement of linguistic stress. Theoretical issues and implications for research and clinical practice are deferred for synthesis of the present findings with those from a multi-site cross-validation project (Shriberg, Aram, & Kwiatkowski, 1997b).

KEY WORDS: apraxia, phonology, speech, children, disorders

The prior paper in this series reviewed descriptive and theoretical perspectives on developmental apraxia of speech (DAS) (Shriberg et al., 1997a). A literature review and a local ascertainment study led to two conclusions. First, the term *suspected DAS* appears to persist clinically as a functional explanation for children with speech delays that in some way differ from error patterns in typical speech delay and that take longer to normalize even with usually sufficient intervention. Second, notwithstanding the lack of a diagnostic marker, the predominant view of DAS is that it is a motor-speech disorder involving deficits in the prearticulatory sequencing of segmental targets. Until a diagnostic marker is validated for this putative clinical entity, clinicians and researchers must acknowledge the circularity in inferences about deficits that define the disorder and the psycholinguistic processes that underlie those deficits.

The primary goal of the two studies reported in this paper is to determine if a diagnostic marker for DAS can be identified in two samples of children with suspected DAS. A secondary goal is to assess the level of support for alternative hypotheses about the nature and origin of DAS, including the possibility that some form of DAS may be genetically transmitted.

Diagnostic Markers and Phenotype Markers

The role of diagnostic markers in clinical and research contexts was discussed in the prior paper (Shriberg et al., 1997a). The validity and reliability of a diagnostic marker is evaluated on its ability to identify persons who do and do not have a disorder or disease. Typical goals for a

 1092-4388/97/4002-0286

candidate diagnostic marker in clinical medicine are sensitivity (true positives) and specificity (true negatives) values of at least 90%.

In addition to the question of a diagnostic marker for DAS, an etiologic issue is whether the same or other characteristics can also serve in genetic studies as the phenotype marker of DAS. The literature supports the likelihood that DAS is an inherited disorder (cf. Shriberg et al., 1997a). As above, however, the phenotypes used in familial studies reflect only features on the many checklists presumed to characterize DAS. Appropriate phenotype markers for behavioral and molecular genetic designs have been discussed by Pennington (1986), who proposes that ideal phenotype markers meet five criteria. First, the marker must be expressed in all genotypes that code for the disorder (complete penetrance), have early onset, and have developmental persistence. Second, although some genetic models permit continuous phenotypes, it is preferable that the distribution of the marker be bimodal, so that family members can be identified as affected or not affected. Third, the marker "must be present in all individuals who meet standard diagnostic criteria and in some but not all of their relatives who do not meet standard diagnostic criteria" (p. 72). Here Pennington draws the important distinction between diagnostic and phenotype markers. Specifically, to be useful for genetic analysis procedures it is the phenotype marker, not the diagnostic marker, that has to differentiate affected versus nonaffected individuals within families. Fourth, the phenotype marker must have a "logical and potentially causal" relationship to the full-blown disorder. That is, the ideal phenotype marker should be specific to the syndrome and not found in other complex behavioral disorders. Finally, following Elston and Namboodiri (1980), Pennington suggests that phenotype markers can be defined by multivariate procedures, but they ideally should involve only a single test or trait. Although the present studies seek to identify a diagnostic marker for DAS, findings will also be discussed from the perspective of identifying a potential phenotype marker that meets Pennington's five criteria.

Study I: Speech and Prosody-Voice Characteristics of Children With Suspected DAS

Method

Procedures

Conversational speech samples and articulation test responses were obtained from 14 children in the Cleveland area who had previously been seen in a study series in developmental apraxia of speech. All assessment sessions were conducted by the second author, who was familiar with the children from the prior research. The protocol for making high quality audiocassette recordings of conversationally rich speech samples from the children has been reported in previous studies of children with normal and disordered speech (e.g., Shriberg, 1986, 1993; Shriberg & Kwiatkowski, 1982, 1994). Speech samples were recorded on a well-maintained, high-quality audiocassette tape recorder with a remote microphone positioned approximately 15 cm from a child's lips. Examiner topics for the conversational speech samples focused on the children's activities and interests. The examiner verbally glossed all utterances that were difficult to understand, thereby demonstrating comprehension to the speaker and providing a context for later orthographic and phonetic transcription.

Subjects

Table 1 provides information for 14 children with suspected developmental apraxia of speech seen at a diagnostic center at Rainbow Babies and Children's Hospital in Cleveland, Ohio, including age, gender, and historic/test data on episodes of otitis media, intelligence, language, and oral-peripheral movements. Conversational samples for the present study were obtained when children were from 4 years 10 months to 14 years 11 months (M = 7 years 11 months, SD = 3 years 1 month). Hearing was normal at the time of assessment as demonstrated by passing a pure tone audiometric screening evaluation at 25 dB bilaterally for 500, 1000, 2000, and 4000 Hz. Each child's oral-peripheral speech mechanism was evaluated using a clinical protocol adapted from Spriestersbach, Morris, and Darley (1978). For the purpose of excluding dysarthric children, lip, tongue, and velar range and precision of movement were assessed and in all cases were judged to be clinically normal. A battery of single and sequential volitional oral movements, also adapted from Spriestersbach et al. and including nonspeech oral movements such as whistling, blowing, and kissing, was administered for the purpose of assessing the presence or absence of oral apraxia (Aram & Horwitz, 1983).

To accommodate potential differences associated with age over this 10-year range, 7 (50%) of the children were placed in a *younger* subgroup (less than 7 years of age). The remaining 7 (50%) children were placed in an *older* subgroup (7–15 years of age). The ages of children in the younger subgroup were comparable to the ages of children typically studied in clinical research in developmental phonological disorders, whereas there are few comparable data available for children within the age range of the older subgroup. A

Table 1. Descriptive information for 14 children with suspected DAS seen at a diagnostic research center in Cleveland, Ohio.

Child	Age (yrs; mos)	Gender	Reported episodes of otitis media	Intelligence[a]	Language comprehension[a]		Language expression[a]		Oral-peripheral nonspeech movements	
					PPVT (yrs;mos)	Syntax	Naming (yrs;mos)	Syntax	Full-range	Oral apraxia
Younger subgroup										
1	4;10	M	>4	Leiter: >109	8;0	ACLC: Normal	Normal	Mild to moderate deficit	Normal	Absent
2	5;0	F	Unknown/ adopted	WPPSI: 80	Moderate deficit reported	ACLC: Normal	Word finding problems	Moderate deficit	Normal	Present
3	5;9	M	None	Leiter: 100	7;5	NSST: Between 50th and 75th percentile	Word finding problems	Severe deficit	Normal	Present
4	5;10	M	1	Leiter: 102	4;11	NSST: Between 50th and 75th percentile	EOWPVT: 3;4	Severe deficit	Normal	Absent
5	6;1	M	None	Normal by report	9;5	Token: Mild memory deficits, normal syntax	EOWPVT: 6;10	Severe deficit	Normal	Absent
6	6;3	M	Recurrent	Leiter: 88	7;0	Token: Mild memory and moderate syntactic deficits	EOWPVT: 4;9	Moderate deficit	Normal	Absent
7	6;3	M	None	WPPSI: Reported normal	7;0	Token: Mild memory and syntactic deficits	EOWPVT: 6;10; word finding problems	Moderate deficit	Normal	Absent
Older subgroup										
8	7;1	F	None	Leiter: 99	7;3	NSST: 90th percentile	EOWPVT: 1 year below CA	Moderate to severe deficit	Normal	Absent
9	7;5	M	None	Leiter: 93	Mild deficit reported	Token: Mild memory deficits, normal syntax	Word finding problems	Moderate to severe deficit	Normal	Present
10	8;6	M	None	WISC-R: 104	8;7	Token: Normal	Word finding problems	Moderate to severe deficit	Normal	Absent
11	9;11	M	Recurrent; PE tubes at 2-1/2 years	Leiter: 76	7;7	Token: Mild memory deficits, normal syntax	EOWPVT: 8;6	Moderate to severe deficit	Normal	Absent
12	10;2	M	None	Leiter: 84 WISC-R: 81	Not given	Token: Mild memory and syntactic deficits	EOWPVT: 9;11; word finding problems	Moderate to severe deficit	Normal	Absent
13	12;11	M	None	Leiter: 66	8;3	Token: normal memory; mild syntactic deficits	Word finding problems	Moderate deficit	Normal	Present
14	14;11	M	None	Leiter: 96	Normal	NSST: Normal	Word finding problems	Moderate deficit	Normal	Present

[a]Abbreviations refer to the following tests: ACLC, Assessment of Children's Language Comprehension (Foster, Giddian, & Stark, 1973); EOWPVT, Expressive One-Word Picture Vocabulary Test (Gardner, 1979); Leiter, Arthur Adaptation of the Leiter International Performance Scale (Arthur, 1952); NSST, Northwestern Syntax Screening Test (Lee, 1964); PPVT, Peabody Picture Vocabulary Test (Dunn, 1965); Token, The Token Test for Children (DiSimoni, 1978); WPPSI, Wechsler Preschool and Primary Scale of Intelligence (Wechsler, 1967); WISC-R, Wechsler Intelligence Scale for Children–Revised (Wechsler, 1974).

total of 12 of the 14 children were males (86%), with the 2 females among the youngest children in each of the two subgroups. As shown in Table 1, the otitis media, intelligence, language, and oral-peripheral status of these children ranged from normal to involved on a variety of measures. As obtained by parental report, all but 2 of the 14 children had relatives with speech, language, or learning disorders, primarily nuclear family members.

Table 2 provides summary speech data for the 14 children with suspected DAS, based on conversational speech samples obtained by the examiner in Cleveland and processed in Madison (see below). Percentage of Consonants Correct (PCC) scores (Shriberg & Kwiatkowski, 1982; Shriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986) ranged from 43.5% to 89.4%, with a mean of 67.0% for all 14 children. Average PCC scores of the children in the older subgroup (77.4%) were over 20 percentage points higher than average PCC scores of children in the younger subgroup (56.7%). Accordingly, the adjectives corresponding to PCC severity scores (unadjusted for the standard error of measurement, language status, and/or prosody-voice status; cf. Shriberg & Kwiatkowski, 1982; Shriberg et al., 1986), reflect less severe segmental involvement for children in the older subgroup.

The rightmost column in Table 2 indicates each child's classification status as determined by the Speech Disorders Classification System (SDCS) (Shriberg, 1993). The classification categories in the SDCS are based on a synthesis of reference data on normal and disordered phonological and phonetic development (cf. Appendix in Shriberg, 1993) and on a revision using an extended database of life span reference data (Shriberg, Austin, Lewis, McSweeny, & Wilson, in press-b). It is useful to describe the abbreviations in the rightmost column of Table 2 as those corresponding to stems, affixes, and brackets.

The stem *NSA* is the abbreviation for *normal (or normalized) speech acquisition*. The stem *NSA/SD* (which was not assigned any of the children in Study I, but will occur in later discussion) is for speech that falls on the borderline between NSA and *speech delay (SD)*

Table 2. Speech status of 14 children with suspected DAS, seen at a diagnostic research center in Cleveland, Ohio.

		Age at assessment		Percentage of Consonants Correct (PCC)		
Child	Gender	Months	Yrs;mos	PCC	Severity level	SDCS[a]
Younger subgroup						
1	M	58	4;10	43.5	Severe	SD[+]
2	F	60	5;0	51.9	Moderate-Severe	SD[+]
3	M	69	5;9	52.9	Moderate-Severe	SD
4	M	70	5;10	44.6	Severe	SD[+]
5	M	73	6;1	62.7	Moderate-Severe	SD[+]
6	M	75	6;3	65.2	Mild-Moderate	SD
7	M	75	6;3	76.1	Mild-Moderate	SD
Subtotal:	*M*	68.6	5;9	56.7		
	SD	6.9	0;7	11.9		
Older subgroup						
8	F	85	7;1	87.2	Mild	NSA–
9	M	89	7;5	51.0	Moderate-Severe	SD[+]
10	M	102	8;6	82.6	Mild-Moderate	NSA-[+]
11	M	119	9;11	68.8	Mild-Moderate	RE-2
12	M	122	10;2	89.4	Mild	[RE-3]
13	M	155	12;11	80.5	Mild-Moderate	RE-2
14	M	179	14;11	82.3	Mild-Moderate	[RE-2][+]
Subtotal:	*M*	121.6	10;2	77.4		
	SD	34.6	2;11	13.4		
Total:	*M*	95.1	7;11	67.0		
	SD	36.5	3;1	16.2		

[a]Speech Disorders Classification System (Shriberg, 1993). See text for description of the classification entries.

JSLHR, Volume 40, 286–312, April 1997

(cf. Shriberg et al., in press-b). The stem *SD* is the classification for children up to 9 years of age who have speech-sound deletions and/or substitutions that are inappropriate for their age. On the assumption that 9 years is the sociobiological end point for the construct of speech *delay* (Shriberg, 1993), the stem *RE* is the abbreviation for several subtypes of *residual errors* in children 9 years of age and older. *RE-1* (not occurring in Table 2) indicates only common clinical distortion errors (e.g., dentalized /s/, derhotacized /r/, velarized /r/), *RE-2* indicates both common clinical distortion errors and continuing imprecise speech (i.e., continuing omissions and deletions of the type seen before 9 years in SD), and *RE-3* indicates imprecise speech alone.

The prefix "*Q*" (not occurring in Table 2) in a child's SDCS classification (e.g., QSD, QRE) indicates that the classification is *Q*uestionable. QSD indicates that the child's error pattern would be considered normal if the child were one year younger; QRE is used for 6- to 8-year-old children with common clinical distortions and some specific substitutions (i.e., all forms of RE are questionable below 9 years of age). The suffix "–" added to NSA stems (see Children 8 and 10) indicates the presence of one or more speech-sound errors that are acceptable for the child's age. The suffix "+" after any stem indicates that the error pattern also includes uncommon clinical distortions in over 20% of the words in the sample (cf. Appendix, Shriberg, 1993).

Finally, the brackets around the classifications indicate marginal placement in this classification category. Bracketed stems meet some, but not all, of the type/token criteria for the stem, and a bracketed "+" indicates uncommon clinical distortions in only 10%–20% of the words in the sample.

Compared to their PCC scores and adjectives, which are in the mild and mild-moderate range for 8 of the 14 children in Table 2, the SDCS classifications provide qualitative support for why these children are suspected to have DAS. In the younger group, all 7 children meet SDCS criteria for SD, with 4 of the 7 also marginally meeting criteria for SD+. For the older children, including 3 children younger than 9 (Children 8-10), all except Child 8 and Child 10 have either SD (Child 9), or nonmarginal or marginal forms of RE that include imprecise speech (Children 11–14). Thus, although PCC scores for most of these older children are in the Mild and Mild-Moderate range, their persisting speech errors include age-inappropriate deletion and/or substitution errors.

Comparison Groups and Data Reduction

Comparison Groups

The analysis to follow uses comparison data from children whose phonological disorders were not suspected to reflect developmental apraxia of speech. One group, termed the *younger speech delayed* group (*younger SD*), includes samples from 64 3- to 6-year-old children (mean age = 4 years 3 months) with developmental phonological disorders of unknown origin. Detailed speech data from these children were presented in Shriberg and Kwiatkowski (1994) and will be presented in a series of comparisons to data for the younger DAS group in the present study. The second comparison group, termed the *older speech delayed* group (*older SD*), consists of nine 7- to 13-year-old children (mean age = 7 years 11 months) with developmental phonological disorders of unknown origin. Each of these children had been seen for speech assessment and/or speech intervention at the University of Wisconsin-Madison Phonology Clinic. These were the only 9 children in a clinical-research database whose records met the following inclusionary and exclusionary criteria: (a) Each child had a speech disorder of unknown origin persisting past the age of 7 years and, (b) there was no indication that the speech delay might be associated with developmental apraxia of speech. All procedures used to transcribe and prosody-voice code the conversational speech samples from the two comparison groups were similar to those used to process speech samples from the 14 children with suspected DAS, as described next.

Transcription and Prosody-Voice Coding

The conversational speech samples forwarded to Madison were first transcribed by consensus by two of the authors (LS, JK) using procedures described in Shriberg (1986), Shriberg and Kent (1995), and Shriberg, Kwiatkowski, and Hoffmann (1984). Detailed notes were taken on both segmental and suprasegmental behaviors. Several years later, when a procedure for prosody-voice coding became available (Shriberg, Kwiatkowski, & Rasmussen, 1990), a research transcriptionist transcribed and prosody-voice coded the 14 conversational speech samples from children with suspected DAS. The transcriptionist was provided only the children's age and gender.

Extensive reliability data for the systems and procedures described above for phonetic transcription have been reported in Shriberg and Lof (1991). Several forms of test and examiner reliability data and detailed validity data for the prosody-voice coding procedure, including acoustic validation data, are presented in Shriberg, Kwiatkowski, Rasmussen, Lof, and Miller (1992). Interjudge agreement figures for the transcriber were obtained for conversational speech samples from four data sets, which included 1,346 utterances and a total of 6,061 words (Shriberg, Austin, Lewis, McSweeny, & Wilson, in press-a). These reliability samples were obtained from 32 3- to 45-year-old speakers, including some children randomly sampled from the present data set, with the majority of samples from younger children and with speech ranging from normal to severely involved. For consonants, the transcriber's average agreement with three other experienced transcribers was 81.9% for narrow phonetic transcription and 88.7% for broad phonetic transcription. For vowels, average interjudge agreement was 79.1% for narrow phonetic transcription and 85.6% for broad phonetic transcription. Reliability of prosody-voice coding based on conversational speech samples from 28 children randomly selected from a large database of children with mild to severe speech delay, including children with suspected DAS, was presented in Shriberg et al. (1992, Table 6). The transcriber's intrajudge reliability for summative

Table 3. Structural characteristics of the conversational speech samples from reference data and from the four groups of children.

	Reference data[a]		Younger		Older	
Variable	Rank	%	DAS (*n* = 7) %	SD[b] (*n* = 64) %	DAS (*n* = 7) %	SD[c] (*n* = 9) %
Intended word form						
CVC	1	31.9	25.0	29.3	25.4	26.7
CV	2	21.1	21.4	21.2	22.9	20.9
2-Syllable	3	13.5	14.5	14.3	14.9	12.1
VC	4	11.9	15.3	13.8	14.2	15.8
V	5	9.1	13.1	9.2	9.4	9.0
C(n)VCn	6	7.0	4.1	5.3	5.9	7.0
VCn	7	2.2	0.9	1.6	1.1	2.7
CnVC	8	1.5	2.7	2.0	2.8	3.0
3+Syllable	9	1.2	1.7	1.3	2.1	1.1
CnV	10	0.4	1.0	1.3	0.9	1.1
Intended consonant						
n	1	12.0	15.1	13.8	14.7	15.6
t	2	11.8	14.3	12.5	12.5	13.3
s	3	6.9	5.1	6.0	7.8	7.3
r	4	6.8	4.7	5.5	5.8	5.2
d	5	6.1	6.3	5.9	5.7	7.4
m	6	6.0	7.8	7.5	6.7	5.1
ð	7	5.4	2.6	4.6	3.1	4.7
k	8.5	5.3	7.0	6.1	5.6	6.6
l	8.5	5.3	6.0	5.4	5.5	4.8
w	10	4.8	5.1	4.6	4.2	5.4
z	11	4.7	1.1	3.8	3.3	3.9
h	12	4.4	4.8	4.0	4.0	3.2
b	13	3.4	3.6	4.3	4.0	3.6
g	14.5	3.2	3.5	3.4	3.2	2.8
p	14.5	3.2	3.8	3.6	3.8	2.8
f	16	2.1	2.4	1.5	3.0	2.0
j	17	1.7	1.6	1.7	2.1	1.6
ŋ	18	1.6	1.5	1.1	1.2	0.8
v	19	1.5	0.9	1.3	1.5	0.8
ʃ	20.5	0.9	0.7	0.9	0.4	1.5
θ	20.5	0.9	0.5	1.3	0.8	0.7
dʒ	22	0.7	0.7	0.7	0.6	0.2
tʃ	23	0.6	0.9	0.6	0.3	0.7
ʒ	24	0.0	0.0	0.0	0.0	0.0

[a]See references in Shriberg (1986), Appendix C, Table 6a.
[b]Shriberg & Kwiatkowski (1994).
[c]From Phonology Project Database, 1994.

JSLHR, Volume 40, 286–312, April 1997

prosody-voice coding ranged from 85% to 99% point-to-point agreement on the seven prosody-voice parameters described later. Interjudge agreement with the first and second authors on these same samples ranged from 74% to 96% on the seven summative prosody-voice parameters.

Representativeness of the Conversational Speech Samples

Analysis of the structural characteristics of the conversational speech samples from each of the four groups was undertaken to determine if the speech samples had representative percentages of intended word forms and intended consonant targets. Table 3 provides comparison data for the younger and older DAS subgroups, the younger and older SD groups, and a reference data set. The reference data in Table 3 are taken from Shriberg et al. (1986, Table 3, p. 144). The rank orders and percentages in the first data column in Table 3 were derived from four studies of the conversational speech of children developing speech at the normal rate (Carterette & Jones, 1974; Hoffmann, 1982; Irwin & Wong, 1983; Mader, 1954), three samples of children with speech delay (Shriberg & Kwiatkowski, 1982, 1983; Shriberg, Kwiatkowski, & Hoffmann, 1984), and one study of normal-speaking adults (Mines, Hanson, & Shoup, 1978). The rank-ordering and magnitude of the percentages for the younger and older DAS subgroups agree well with these reference data and with data for the respective younger and older SD comparison groups. Linguistic factors underlying the high stability of intended word forms and phonemes in conversational speech samples from diverse subject groups are discussed in Shriberg and Kwiatkowski (1983). For the present purposes, the data in Table 3 support the representativeness of the conversational speech samples for the group and individual data analyses.

Results

Do Children With Suspected DAS Differ From Children With SD on Severity of Speech or Prosody-Voice Involvement?

Speech Severity Indices

The first question posed of the grouped data was whether the 14 children with suspected DAS differed in severity of involvement from children in the two comparison groups. Table 4 includes summary data comparing children with suspected DAS to children with SD on five alternative indices of severity of speech involvement. Several statistical considerations (small samples sizes, disproportionate sample sizes for the younger group comparisons, disproportionate standard deviations) supported the use of nonparametric Wilcoxon-Mann-Whitney tests (Siegel & Castellan, 1988) to test for statistically significant differences between children with suspected DAS and the comparison groups.

Percentage of Consonants Correct (PCC). The first three rows in Table 4 compare PCC data for the children with suspected DAS and children with SD, including the total PCC and sub-indices reflecting articulation of singletons (PCC: Singletons) and clusters (PCC: Clusters). The mean severity scores are descriptively similar for the groups with suspected DAS and with SD, and the rank-order statistics are statistically nonsignificant for five of the six comparisons. The one significant ($p < .05$) difference was that the younger children with suspected DAS averaged over 11 percentage points lower on the PCC: Clusters than the younger children with SD (mean DAS: 38.9%; mean SD: 50.3%). Detailed information on error targets and error types are reviewed in a subsequent section. As indicated in the rightmost column in Table 4, there were no statistically significant differences between the older DAS and comparison SD group on any of the five severity measures.

Percentage of Consonants Correct–Adjusted (PCC-A). The second three rows in Table 4 provide descriptive data and results of inferential statistical tests for the *Percentage of Consonants Correct–Adjusted (PCC-A)*. As described in Shriberg et al. (in press-a), whereas the PCC scores all clinical speech-sound distortions as incorrect, the PCC-A considers certain residual distortion errors (derhotacized consonant and vocalic /r/, dentalized fricatives/affricates, velarized and labialized /l/) as correct. Because the PCC-A "adjusts" a child's or adult's PCC score to reflect such common distortions, the PCC-A is more sensitive to all other types of speech-sound errors. As shown in Table 4, this adjustment yielded two changes from PCC findings. First, it increased the difference in means between the younger DAS and SD groups on clusters (mean DAS: 42.1%; mean SD: 60.6%) that met significance criteria at the .01 level. Second, it increased the difference in means for the total PCC-A scores (mean DAS: 58.6%; mean SD: 69.3%), yielding a second statistically significant difference between the younger DAS group and the comparison group.

Percentage of Consonants Correct–Revised (PCC-R). The next three data rows in Table 4 describe findings and results of inferential statistical tests for the *Percentage of Consonants Correct–Revised (PCC-R)* scores of children in the four groups. The PCC-R considers *all* speech-sound distortions as correct, not just the residual distortions as in the PCC-A. Thus, as a severity metric, the PCC-R is maximally sensitive to potential differences in the percentages of deletion and substitution errors in a speech sample. As shown in Table 4, the average scores of younger children with suspected DAS

Table 4. Severity of involvement of younger and older children with suspected DAS and age-matched comparison groups of children with speech delay (SD).

	Younger					Older				
	DAS (n = 7)		SD (n = 64)			DAS (n = 7)		SD (n = 9)		
Severity metric	*M*	*SD*	*M*	*SD*	W[a]	*M*	*SD*	*M*	*SD*	W
Percentage of Consonants Correct (PCC)										
Singletons	61.4	13.0	66.3	8.5	201.0 *ns*	80.4	11.0	83.1	6.6	58.0 *ns*
Clusters	38.9	8.7	50.3	12.8	131.0 *	68.9	20.5	71.6	11.2	60.5 *ns*
Total	56.7	11.9	62.7	8.2	176.5 *ns*	77.4	13.4	80.0	7.2	58.0 *ns*
Percentage of Consonants Correct–Adjusted (PCC-A)										
Singletons	62.9	12.9	71.7	9.7	161.0 *ns*	83.1	12.1	88.9	4.7	51.5 *ns*
Clusters	42.1	9.8	60.6	16.2	109.0 †	75.9	22.1	80.7	9.9	61.0 *ns*
Total	58.6	12.0	69.3	10.2	134.0 *	81.2	14.5	86.8	5.7	56.5 *ns*
Percentage of Consonants Correct–Revised (PCC-R)										
Singletons	67.1	12.1	74.5	9.2	172.0 *ns*	86.2	12.3	89.6	4.6	57.5 *ns*
Clusters	44.8	9.7	64.5	16.1	99.5 †	77.8	21.9	81.8	9.0	63.0 *ns*
Total	62.5	11.2	72.4	9.8	140.0 *	84.0	14.7	87.5	5.4	59.5 *ns*
Percentage of Vowels Correct (PVC)										
	82.1	3.7	91.4	3.6	46.0 ††	85.5	7.3	92.5	4.0	43.0 *ns*
Intelligibility Index (II)										
	77.3	18.6	91.7	7.9	115.5 †	89.4	13.1	97.5	2.1	43.0 *ns*

[a]Wilcoxon-Mann-Whitney (Siegal & Castellan, 1988)
*$p < .05$
†$p < .01$
††$p < .001$

were again significantly lower than comparison children on the PCC-R for clusters (mean DAS: 44.8%; mean SD 64.5%; $p < .01$) and on the total PCC-R (mean DAS: 62.5%; mean SD: 72.4%; $p < .05$).

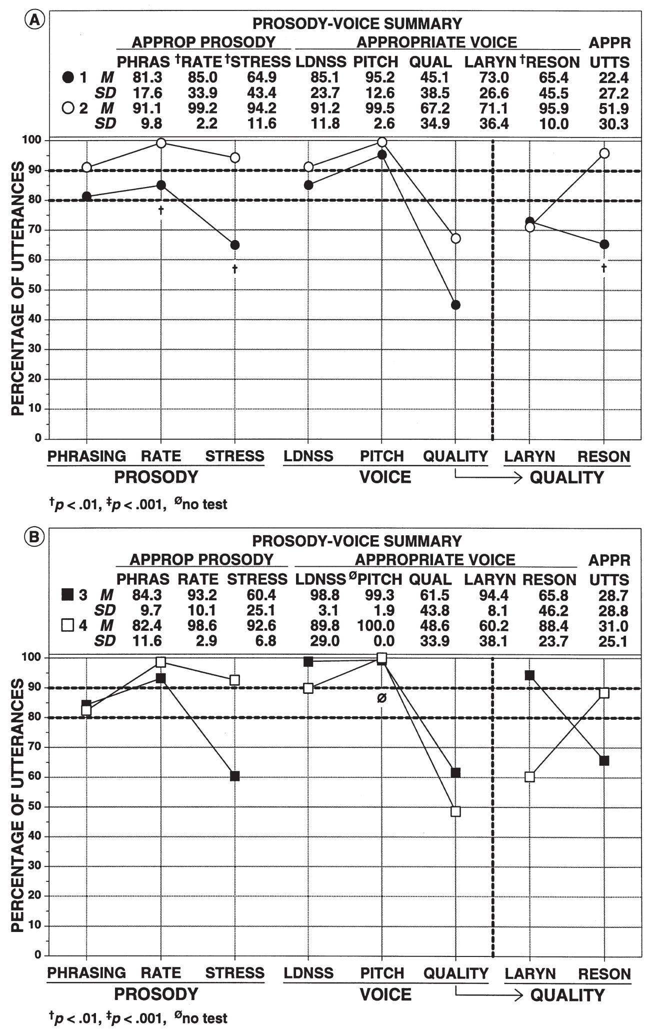
Percentage of Vowels/Diphthongs Correct (PVC). The severity metric assessing children's vowel and diphthong accuracy was the *Percentage of Vowels/Diphthongs Correct (PVC)*. The PVC is obtained from a conversational speech sample using the same percentaging procedures as those used in the PCC metric (Shriberg, 1986), with denominators including all intended vowel and diphthong sounds. As shown, the younger children with suspected DAS had significantly ($p < .001$) lower vowel/diphthong accuracy in conversational speech, averaging approximately 9 percentage points lower than children in the younger comparison group with SD (mean DAS: 82.1%; mean SD: 91.4%).

Intelligibility Index (II). The *Intelligibility Index (II)* reflects the percentage of child-intended words in the conversational speech sample that the transcriber was able to gloss (Shriberg & Kwiatkowski, 1982; Weston & Shriberg, 1992). This metric reflects a best case estimate of a speaker's intelligibility, because the original examiner provided a verbal gloss on the tape and the transcribers were free to use multiple playbacks of difficult strings (cf. Kwiatkowski & Shriberg, 1993). The younger children with suspected DAS were significantly ($p < .01$) less intelligible than the younger children with SD, averaging approximately 14 points lower (DAS: 77.3%; SD: 91.7%).

Prosody-Voice Indices

Description. Figure 1 includes the summary panels from a prosody-voice analysis procedure termed the *Prosody-Voice Screening Profile (PVSP)* (Shriberg, 1993; Shriberg et al., 1990, 1992). This perceptual procedure provides summary and detailed data on seven prosody and voice variables coded from the conversational speech samples. The top panel (Panel A) includes prosody-voice information for the younger DAS group (Group 1: filled circles) and the comparison SD group (Group 2: open circles); the bottom panel (Panel B) includes prosody-voice information for the older DAS group (Group 3: filled squares) and the comparison SD group (Group 4: open squares). Prosody-voice coding of the conversational speech samples was possible for only 62 of the original 64 younger SD children. The arrangement of the data in each of the panels provides summary means and

Figure 1. *Prosody-Voice Profile* comparison of Study I children with suspected developmental apraxia of speech (DAS) and children with speech delay (SD). Panel A data are for the younger children in each group (Groups 1 and 2, respectively) and Panel B data are for the older children in each group (Groups 3 and 4, respectively). Only the data indicating percentages of utterances with appropriate prosody-voice are shown (see text).



standard deviations (numeric section at the top of each panel) and graphic data on the seven suprasegmentals scored in the prosody-voice procedure: *Phrasing, Rate, Stress, Loudness, Pitch, Laryngeal Quality*, and *Resonance Quality*. Notice that Laryngeal and Resonance Quality are also displayed as one combined category. The data points in the graphic section are the percentage of utterances considered *appropriate*, with the two horizontal dashed lines indicating the 90% screening cutoff for *pass* and the 80% cutoff for *questionable pass* (cf. Shriberg et al., 1990, 1992).

The daggers and double daggers in the numeric and graphic sections of each speech profile in Figure 1 and in subsequent figures indicate significant between-group differences at the .01 and .001 levels, respectively. As in Table 4, the inferential statistic was the nonparametric Wilcoxon-Mann-Whitney test (Siegel & Castellan, 1988). Although means and standard deviations provide the most meaningful descriptive statistics for the numeric and graphic displays, nonparametric statistics provide the most appropriate inferential statistical tests. Nonparametric tests are appropriate when the data include (a) small or disproportionate sample sizes (b) non-normal distributions (i.e., relative to skew, kurtosis, and nonequal variances), and (c) correlated means and standard deviations at extremes of measurement. All three constraints are present in the data of Figure 1 and in subsequent speech profile analyses. The two probability levels, .01 and .001, bracket, respectively, liberal and conservative family-wise alpha levels for the combined number of tests in these panels, and additional comparisons on more detailed levels of the output not shown here (i.e., the .05 alpha level is not appropriate). These statistical approaches to prosody-voice (and later speech profile) analysis attempt to balance the goals of exploratory data analysis, advisory inferential tests, and avoidance of Type I or Type II errors of inference.

Younger children with suspected DAS. As shown in the numeric and graphic sections of Figure 1, Panel A, younger children with suspected DAS had lower average percentages of appropriate utterances on each of the seven suprasegmentals except laryngeal quality. Statistically significant between-group differences at the .01 alpha level were obtained for rate (DAS: 85.0%, SD: 99.2%), stress (DAS: 64.9%, SD: 94.2%) and resonance quality (DAS: 65.4%, SD: 95.9%).

Analyses in additional panels of the prosody-voice profiles (not shown here) indicated that the statistically significant differences in Figure 1, Panel A, were primarily associated with three types of inappropriate prosody-voice codes. The rate differences were primarily due to the DAS children averaging more utterances considered inappropriate on a PVSP code termed *Slow Articulation/Pause Time* (Shriberg et al., 1990). The criterion for this code is a rate of fewer than two syllables per second, associated with excessive pause time between words and/or excessively long articulation time. The significant stress differences were primarily associated with a PVSP code termed *Excessive/Equal/Misplaced Stress*. This code was used when children's utterance included either (a) a monostressed pattern characterized by forceful, punctuated stress; (b) misplaced word stress relative to expected phrasal or emphatic stress patterns; or (c) blocks or sound prolongations (Shriberg et al., 1990). Finally, the significant resonance quality differences were associated with the DAS children averaging more utterances coded as *Denasal*, used for utterances in which there was a lack of normal nasal resonance on vowels or diphthongs (including contexts in which assimilative nasality would be appropriate). Additional discussion of these groupwise prosody-voice differences is deferred to the second series of analyses concerned with individual children with suspected DAS.

Older children with suspected DAS. As shown in Figure 1, Panel B, there were no statistically significant differences at the $p < .01$ level or less in the prosody-voice scores of older children with suspected DAS compared to children in the older SD group. The trends were somewhat similar to those observed for the younger children, with the older children with DAS having notably lower average scores than the children with SD on stress and resonance quality. Additional discussion of these group-level, prosody-voice findings is again deferred to the individual-level analyses to follow.

Individual Subject Analysis

The findings in Table 4 and Figure 1 indicated that on some severity metrics, for the younger group comparisons in particular, children with suspected DAS scored significantly lower than comparison children with SD. To assess whether an individual's score on any of the speech or prosody-voice severity variables might discriminate group membership, scatter plots were constructed for each measure, even for measures on which statistically significant findings were not replicated in both younger and older group comparisons. Each of the scatter plots was examined to determine the degree to which the measure divided children into two groups based on their severity scores, the requirement of a diagnostic marker.

Beginning with the speech severity indices, examination of each plot indicated that the overlap in subject scores between the two groups was too great to claim use as a potential diagnostic marker. For example, of particular interest were the data from the PVC measure, in which the younger children with suspected DAS scored significantly lower as a group than children with SD. Only 2 of the 7 children with suspected DAS had PVC scores below the lowest PVC score of all children with SD, and these two scores were only marginally lower. Scores for the other 5 younger children with suspected DAS were within the range of scores for the younger children with SD. For the older children with DAS, plots indicated that 4 of the 7 PVC scores were below the lowest PVC score for the children with SD, with the other three scores scattered within the range of scores for the children with SD (including two scores for children with suspected DAS near the top of that range). Similar failures to discriminate children with suspected DAS from those with SD occurred on each of the other speech measures. As described for the PVC analyses, the plotted scores overlapped considerably, with many younger and older children with SD having as severe an involvement on each of the severity measures as children with suspected DAS. Added to the finding that none of the speech severity measures replicated significant findings for both younger and older comparisons, these analyses suggested that severity of involvement was not a viable candidate for a diagnostic marker of DAS.

In contrast to the speech findings, scatter plots for the prosody-voice findings suggested the possibility that stress and quality might be candidates for diagnostic markers. This conclusion was based on both the scatter plot findings and on the replication across age groups. Detailed analyses of subject-level prosody-voice data are deferred until after the following descriptions of additional speech analyses.

Do Children With Suspected DAS Differ From Children With SD on Speech Error Targets or Error Types?

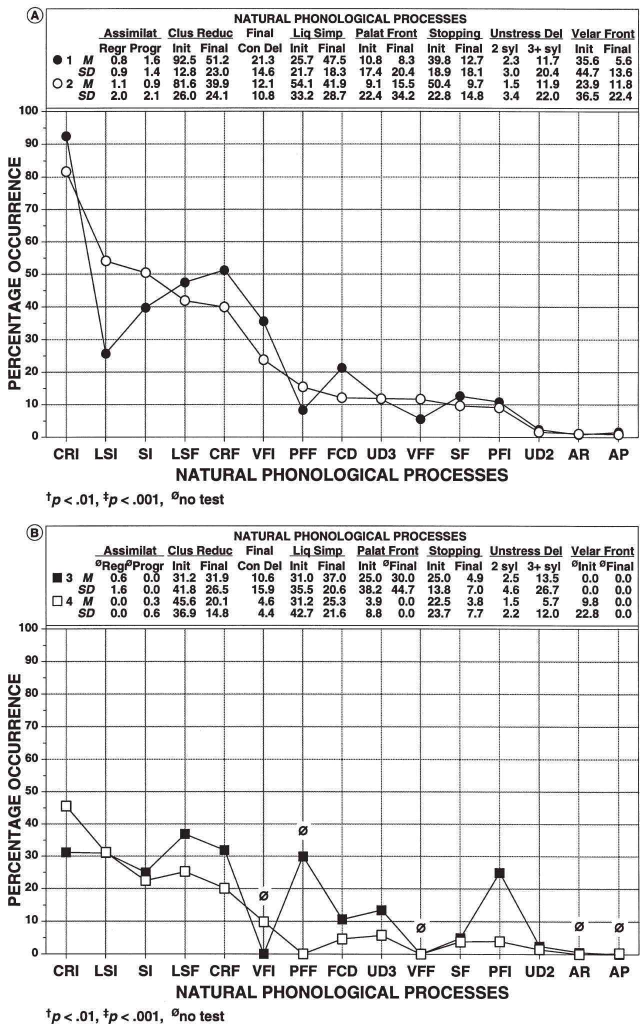
Natural Phonological Processes Analyses

Rationale and procedure. If severity of speech-sound errors does not individually differentiate children with suspected DAS from those with speech delay, what support can be marshaled for specific error targets or error types as a diagnostic feature? One widely used approach to error target/type analysis in child phonology is to profile speech errors by *phonological processes*. Phonological processes reflect speech errors aggregated over Target Sounds x Structural Contexts or Word Forms x Error Types. Thus, processes attempt to account for commonalities among speech-sound errors by positing similar effects at more abstract phonological levels. The value of phonological process analysis as a descriptive device is purported to be its ability to capture generalizations not apparent by analyses using context-free and error-free analytic units, such as phonemes and features. For example, the phonetic motivation for an assimilatory sound change (such as “gog” for “dog”) is presumedly missed by traditional substitution analysis, which captures only the d/g substitution.

Figure 2 is a summary of findings for a phonological process analysis approach in which eight phonological processes meeting criteria for *naturalness* (cf. Shriberg, 1983, 1986, 1991; Shriberg & Kwiatkowski, 1980) are subdivided into 15 categories sensitive to word position (Initial, Final) and other structural elements. The format of the two panels is essentially similar to the format used for the prosody-voice analysis in Figure 1. The top panel (Panel A) includes natural phonological process information for the younger DAS group (Group 1: filled circles) and the comparison SD group

JSLHR, Volume 40, 286–312, April 1997

Figure 2. Natural phonological process analysis comparisons of Study I children with suspected DAS and SD (see Figure 1 for additional panel identification). The abbreviations for processes as arranged left-to-right in the graphic section are as follows: CRI: Cluster Reduction-Initial; LSI: Liquid Simplification-Initial; SI: Stopping-Initial; LSF: Liquid Simplification-Final; CRF: Cluster Reduction-Final; VFI: Velar Fronting-Initial; PFF: Palatal Fronting-Final; FCD: Final Consonant Deletion; UD3: Unstressed Syllable Deletion—3 or more syllables; VFF: Velar Fronting-Final; SF: Stopping-Final; PFI: Palatal Fronting-Initial; UD2: Unstressed Syllable Deletion—2 syllables; AR: Assimilation-Regressive; AP: Assimilation-Progressive.



(Group 2: open circles); the bottom panel (Panel B) includes phonological process information for the older DAS group (Group 3: filled squares) and the comparison SD group (Group 4: open squares). The numeric section at the top of each panel in Figure 2 includes means and standard deviations for the eight natural processes and their subtypes arranged in alphabetical order. The graphic section provides for visual examination of the means, arranged left-to-right in decreasing order based on data obtained in prior studies (Shriberg & Kwiatkowski, 1994). Wilcoxon-Mann-Whitney nonparametric tests at the two advisory alpha levels were obtained, using the same psychometric rationale as discussed above for the prosody-voice analysis.

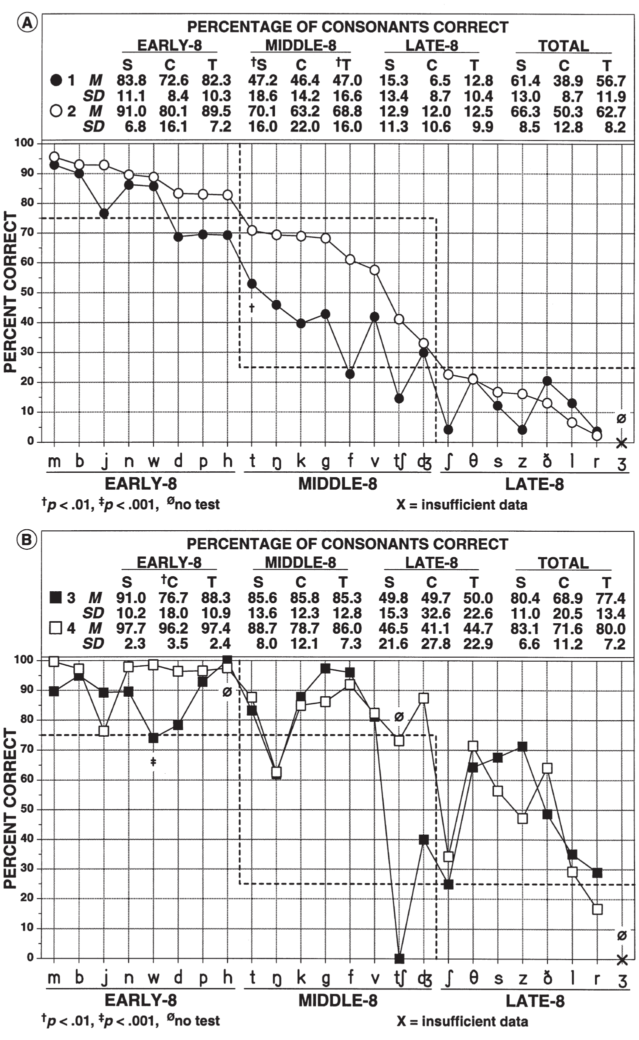
Findings. Of the 30 between-group statistical comparisons for the data in the two panels in Figure 2, none reached significance at the .01 level. The interleaved trends for both the younger and older comparison groups, and the generally dissimilar patterns across younger and older groups with DAS, indicate that children with suspected DAS had essentially the same error targets and error types as children with SD.

Thus, data in Figure 2 suggest that errors aggregated by natural phonological processes do not differentiate children with suspected DAS from children with SD. Additional analyses comparing the number of uncoded substitution and deletion errors in each group (e.g., initial consonant deletion) also indicated no statistically significant differences between children with suspected DAS and children with SD. Perhaps reliable differences in the speech of children with DAS compared to children with SD are only apparent at "higher" (i.e., more linguistically abstract) or "lower" (i.e., more phonetically detailed) levels of the data. The following analyses were undertaken to address potential error target/type differences at phonemic and subphonemic levels, using what has been termed traditional *substitution* or *relational* analyses (cf. Stoel-Gammon & Dunn, 1985).

Speech Profile Analysis: Rationale and Procedure

The two panels in Figure 3 are based on a procedure for profiling speech from conversational samples described in Shriberg (1993). In each panel, the numeric section at the top and the larger graphic section below provide information on the speech status of children with suspected DAS compared to children with SD. As in Figures 1 and 2, the top panel (Panel A) includes speech information for the younger DAS group (Group 1: filled circles) and the comparison SD group (Group 2: open circles); the bottom panel (Panel B) includes speech information for the older DAS group (Group 3: filled squares) and the comparison SD group (Group 4: open squares). The consonant sounds in each of the panels in Figure 3 are divided into groups termed the *Early-8* sounds, the *Middle-8* sounds, and the *Late-8* sounds. Division of the 24 English consonants into these three *developmental sound classes* was based primarily on breaks in the monotonic trend for a large group of 3- to 6-year-old children with speech delay (Shriberg, 1993; the younger SD group in this study). As indicated by the areas demarcated by the dashed lines in the graphic section, scores for the Early-8 sounds in the reference children with SD average above 75% correct, for the

Figure 3. *Speech Profile: Consonants* comparison of Study I children with suspected DAS and SD (see Figure 1 for additional panel identification and see text).



Middle-8 sounds average from 25–75% correct, and for the Late-8 sounds average from 0–25% correct.

The numeric section at the top of Panel A provides mean and standard deviation data for consonant singletons (*S*), consonant clusters (*C*), and all consonants (*T*) for each of the three developmental sound classes and across all 24 sounds (*Total*). Additional panels and statistical output not shown in Figure 3 provided detailed information on error types, including information on 45 diacritic-level allophone differences and speech-sound distortions. These analyses provided the proportion of error types classified as omissions, substitutions, and distortions. The error-type percentages for each target sound and developmental sound class were computed by using each child's number of errors as the denominator, with numerators indicating the number of errors that were omissions, substitutions, or distortions. Thus, unlike the process analyses above, the group-level error type/target analyses to follow are independent of severity of involvement (i.e., each calculation used each child's total number of errors as the denominator, rather than total number of intended sounds). The rationale and procedures for all descriptive and inferential statistical analyses are the same as those described above for the prosody-voice analysis (Figure 1).

Speech Profiles: Consonant Targets

Younger children with suspected DAS. Beginning with findings for Figure 3, Panel A, two observations suggest that the pattern of consonant error targets of younger children with DAS cannot be used to differentiate children with this suspected disorder from children with SD. First, as best shown in the graphic section of Panel A, the left-to-right decreasing slopes of the trends for each group are quite similar. As indicated by the subtotals (*T*) for the three developmental sound classes in the numeric section of this panel, both groups of children have the highest average percentage correct on the Early-8 sounds (DAS: 82.3%, SD: 89.5%), the next highest average percentage correct on the Middle-8 sounds (DAS: 47%, SD: 68.8%), and lowest average percentage correct on the Late-8 sounds (DAS: 12.8%, SD: 12.5%). Notice that, although the children with DAS are generally more involved across all sound classes, the only comparisons on which these severity differences reach statistical significance at the .01 level or greater are within the Middle-8 sound class, for which statistically significant differences within the numeric section were obtained for Middle-8 Singletons (*S*) (DAS: 47.2%, SD: 70.1%, $p < .01$) and Middle-8 Total (*T*) (DAS: 47%, SD: 68.8%, $p < .01$). Only one of the 23 sound-level comparisons computed in the graphic section reached significance: Children with suspected DAS had significantly lower percentage correct scores for /t/ (DAS: 52.9%, SD: 70.9%, $p < .01$).

Speech profile trends similar to those shown in Figure 3 were also obtained for the five manner features: nasals, glides, stops, affricates, and liquids. The younger children with DAS had significantly lower average scores for stops (DAS: 58.9%, SD: 76.9%, $p < .01$). The increased sensitivity of Middle-8 scores on many variables of interest in child phonology has been discussed elsewhere (e.g., Shriberg, Gruber, & Kwiatkowski, 1994). Essentially, because all children may have relatively little difficulty articulating Early-8 sounds correctly and considerable difficulty articulating Late-8 sounds correctly, computations on overall severity metrics such as the PCC can obscure real between-group differences at lower levels of the data. Thus, as visibly apparent in the graphic section of the younger children comparisons in Figure 3, and as documented by the means data in the numeric section, performance of the two groups is most divergent in the middle section of the descending trend reflecting consonant mastery of the Middle-8 sounds.

Using speech profile output similar to the error-*target* analyses in Figure 3, error-*type* analyses were performed comparing the error-type patterns of younger children with suspected DAS to those obtained for children in the younger comparison group. Children with suspected DAS had proportionally more omission errors than children with SD. There were significant differences in the total proportion of omission errors across all consonant sounds (DAS: 42.1%, SD: 25.4%, $p < .01$), for the class of Late-8 sounds (DAS: 38.1%, SD: 17.0%, $p < .01$), and individually for the Middle-8 sound /f/ (DAS: 31.6%, SD: 8.7%, $p < .01$) and the Late-8 sound /z/ (DAS: 72.0%, SD: 13.0%, $p < .001$). At the feature level, the fricative errors of younger children with suspected DAS were significantly more often omissions (DAS: 43.0%, SD: 16.4%, $p < .01$) and significantly less often distortions (DAS: 11.0%, SD: 34.5%, $p < .01$). These data are also consistent with patterns observed in children with both normal and delayed speech development, wherein the error-hierarchy is from omissions, to substitutions, to distortions. That is, because the younger group of children with suspected DAS was somewhat more severely involved than the comparison children with SD, their error-types followed the developmental expectation of proportionally more omission-type errors, particularly on the presumably more difficult sounds to master (i.e., five of the Late-8 sounds are fricatives). Additional information on the specific types of substitutions and distortions is deferred to later analysis at the level of individual subjects.

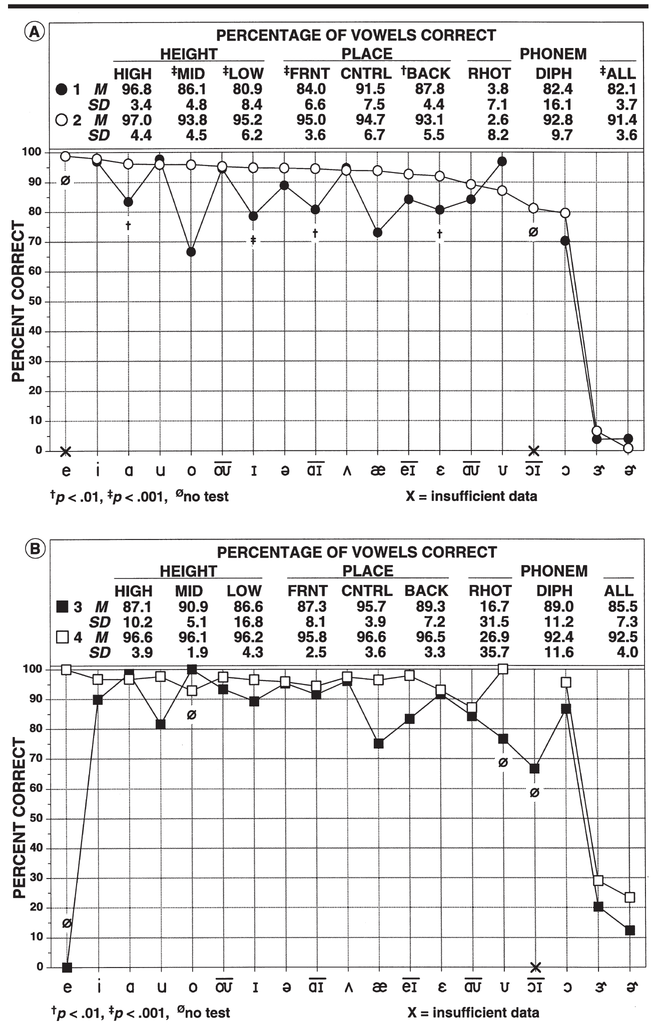
Older children with suspected DAS. The graphic and numeric data in Figure 3, Panel B, also indicate similar trends in consonant mastery for the older children with suspected DAS compared to the group of older children with SD. There is one statistically significant difference in the numeric panel indicating that the older children with suspected DAS had significantly lower percentages of correct clusters (*C*) involving Early-8 sounds (DAS: 76.7%, SD 96.2%, $p < .01$). The only sound-level difference is that the older children with suspected DAS had significantly lower mastery of the Early-8 sound /w/ (DAS: 74.0%, SD: 98.5%, $p < .001$). Findings from the features profile supported the latter findings, with older children with suspected DAS having significantly lower scores on glides (DAS: 78.6%, SD: 97.3%, $p < .001$).

The error-type analyses for the older groups indicated one statistically significant between-group difference in the proportion of error types. The glide errors of older children with suspected DAS were significantly more frequently distortions (DAS: 83.3%, SD: 6.7%, $p < .01$). Calculated as absolute rather than relative errors, the older children with suspected DAS averaged 16.6% distortion-type errors on glides, whereas older children with SD averaged 0.2% glide distortions ($p < .001$). Within the two glides, the distortion errors were distributed across both /w/ and /j/. Additional information on error types is deferred to the second stage analysis of individual subjects.

Speech Profiles: Vowel/Diphthong Targets

The two panels in Figure 4 provide summary information on the vowel/diphthong articulation of the younger (Panel A) and older (Panel B) children with suspected DAS in comparison to children with delayed speech. The analysis formats and statistical conventions for these vowel/diphthong data are similar to those described for Figures 1–3. In the numeric section of each

Figure 4. *Speech Profile: Vowels-Diphthongs* comparison of Study I children with suspected DAS and SD (see Figure 1 for additional panel identification and see text).



panel, descriptive statistics are aggregated separately for sounds classed by Height (High, Mid, Low), Place (Front, Central, Back), the Rhotic vowels (/ɝ/, /ɚ/), the Phonemic Diphthongs (/a͡ɪ/, /a͡ʊ/, /ɔ͡ɪ/), and across all vowels and diphthongs.

Younger children with suspected DAS. Starting with Panel A, several statistically significant differences in both the numeric and graphic sections indicated that the younger children with suspected DAS had lower average mastery of vowels and diphthongs. As also shown in Table 4, younger children with suspected DAS scored significantly lower (by approximately 9%) across all vowel-diphthong sounds (DAS: 82.1%, SD: 91.4%, $p < .001$). Significantly lower percentages were also obtained for vowel/diphthong comparisons summed for the Mid sounds (DAS: 86.1%, SD: 93.8%, $p < .001$), Low sounds (DAS: 80.9%, SD: 95.2%, $p < .001$), Front sounds (DAS: 84.0%, SD: 95.0, $p < .001$), and Back sounds (DAS: 87.8%, SD: 93.1%, $p < .01$). Also, as shown in the top panel of Figure 4, the younger children with DAS had significantly lower average percentages correct for 4 of the 17 individual vowel/diphthong contrasts that could be tested: /ɑ/, /ɪ/, /a͡ɪ/, and /ɛ/.

Separate error-type analysis profiles were obtained to inspect the proportion of substitutions and distortions (omissions were treated as deleted syllables) classified as clinical and nonclinical vowel/diphthong errors (cf. Shriberg, 1993, Appendix). There were no statistically significant differences at the .01 level or less in the distribution of error types on vowels/diphthongs. The only significant finding for this analysis was that the two younger groups differed significantly on the percentages of vowel/diphthong sound changes that were classified as nonclinical differences (i.e., vowel/diphthong changes that plausibly could have been conditioned by the phonetic context). Such differences occurred for 6.3% of the vowel/diphthong sounds of the younger children with suspected DAS, and 10.9% of the vowel/diphthong sounds of younger children with SD ($p < .01$). Further inspection of the specific types of distortion errors made by the younger children with suspected DAS is deferred for a later series of individual analyses.

Older children with suspected DAS. As shown in Figure 4, Panel B, there were no statistically significant differences between the vowel/diphthong mastery comparisons of the older children with suspected DAS and those with SD. Inspection of the error-type data indicated no significant differences in the proportion of errors classified as substitutions or distortions. The two older groups differed significantly on the percentages of vowel/diphthong sound changes classified as nonclinical differences for the central vowel sounds (DAS: 5.3%, SD: 1.4%, $p < .01$). This is in the opposite direction of the differences found for the younger group comparison.

Speech Profiles: Consistency Analysis

As reviewed previously, a cardinal characteristic of children with suspected DAS is reduced token-to-token consistency of errors, especially on repeated tokens of the same word. For the present data, a program was written to identify the percentage of consistent errors on all words occurring in the conversational speech samples. The Speech Profiles framework was used to provide consistency information at the levels of speech sounds, developmental consonant sound classes (i.e., Early-8, Middle-8, Late-8), consonant and vowel features, and overall totals.

Procedure. For each target speech sound, the program identified all word types that occurred at least twice in the sample and in which a target speech sound was produced incorrectly at least twice. A word *type* was defined by the intended target form (e.g., *car* and *cars* are two different word types). An *error class* was defined at two levels: the phoneme level and the phoneme-diacritic level. At the phoneme level, an error class was defined as a deletion or specific substitution for a sound; diacritic level differences (distortion errors and/or modifications of a substituted sound) were disregarded. For example, at the phoneme level, [d/ð] and [t/ð] represent two error classes, whereas [d̥/ð] and [d/ð] are included in the same error class (i.e., the devoiced distortion is ignored). At the phoneme-diacritic level, an error class distinguished diacritic-level differences in substitutions. At this level, the [d̥/ð] and [d/ð] would represent two different error classes. Per-child error consistency percentages were then computed using the following formula:

$$\text{Error Consistency} = \frac{\Sigma \text{ Most Frequent Error Class per Word Type} - \Sigma \text{ Word Types}}{\Sigma \text{ Tokens per Word Type} - \Sigma \text{ Word Types}} \times 100$$

A specific example at the phoneme level may be helpful. Consider a brief speech sample, in which a child says *dog* three times:

I play with do_
My dod
Take dod out

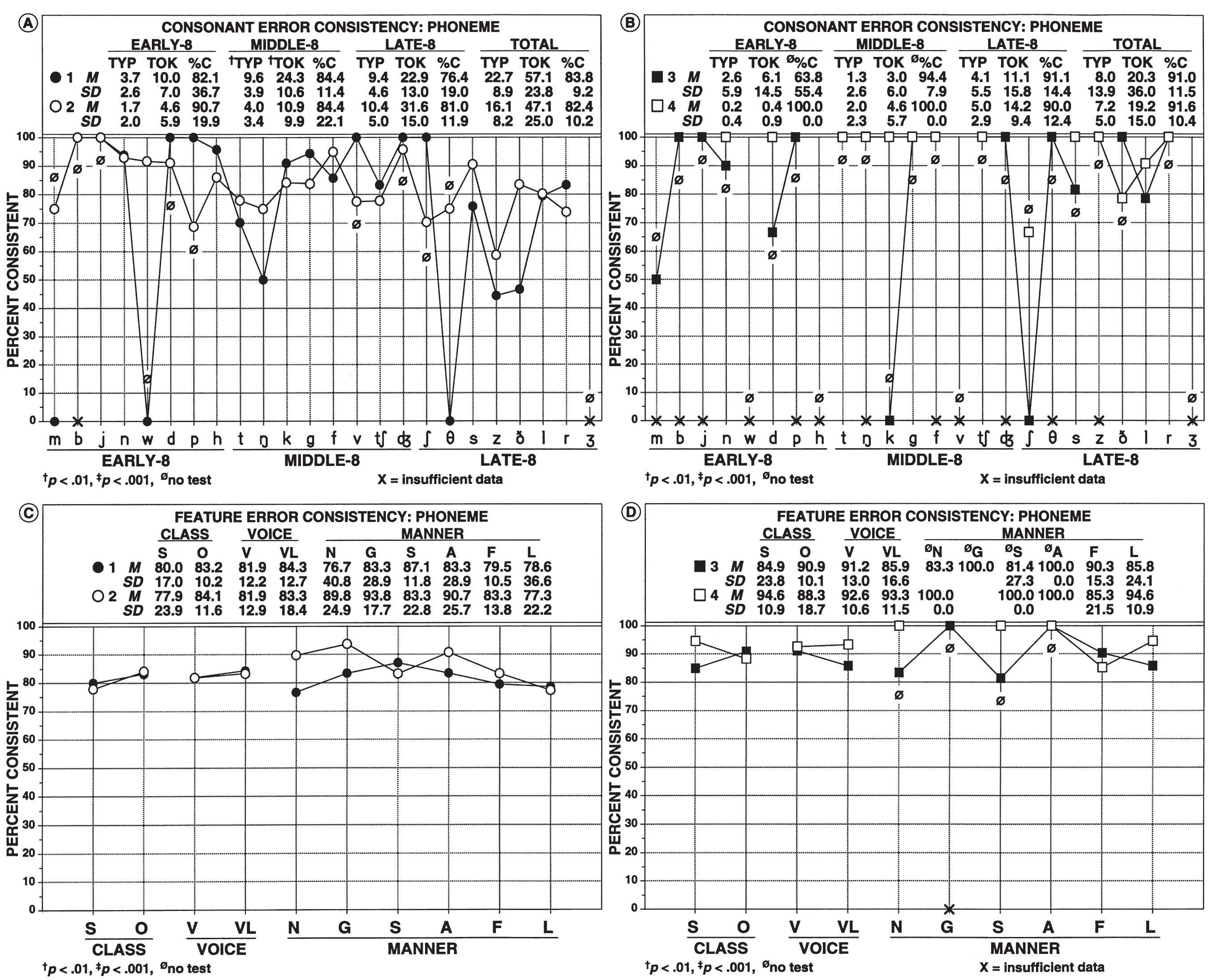
There is one word type (dog), three tokens of the word type, two occurrences of the error class d/g and one occurrence of the error class _/g (omission of g). The d/g substitution is the "most frequent error class per word type." Assuming these are the only word types that include the target sound /g/ in the sample, error consistency would be calculated as follows:

$$\begin{aligned}\text{Error Consistency} &= \frac{2\ (\Sigma \text{ d/g substitutions}) - 1\ (\Sigma \text{ Word Types})}{3\ (\text{tokens of dog}) - 1\ (\Sigma \text{ Word Types})} \times 100 \\ &= \frac{(2-1)}{(3-1)} \times 100 \\ &= 50\%\end{aligned}$$

This percentage reflects *consistency* of errors, rather than *inconsistency* or diversity of errors. Although some type of inconsistency or diversity metric would be sensitive to the number of different error classes per word type, its reliability would be questionable (reflecting as few as one token of an error class), and it could not be expressed as a percentage.

Figure 5 includes consonant error consistency data for phonemes and features. The format of the panels differs from previous formats. The top panels are the consonant error consistency data by phonemes for the younger (Panel A) and older (Panel B) comparisons. The bottom panels are the consonant error consistency data by features for the younger (Panel C) and older (Panel D) comparisons. As with the previous speech profiles, group-level error consistency percentage data are averaged across members of each group. The numeric panels for the phoneme comparisons (Panels A and B) include means and standard deviations for the number of word types (TYP) and the number of tokens per type (TOK) used to calculate the percentage of consistent errors (%C). Any significant between-group difference in the total number of types and/or tokens would suggest a sampling constraint on significant findings for error consistency. As indicated in the numeric section of Panel A, there were two such constraints for Middle-8 sounds. Significantly more types and tokens were used to calculate the consonant error consistency of the younger group with suspected DAS, compared to the younger children with SD.

Figure 5. Consonant error consistency comparison of Study I children with suspected DAS and SD. Groups 1 and 2 are the younger children with suspected DAS and SD, respectively, and Groups 3 and 4 are the older children with suspected DAS and SD, respectively. Panels A and B are the error consistency data for consonant phonemes and Panels C and D are the error consistency data by consonant features (see text).



As indicated in all four panels in Figure 5, there were no significant between-group differences in error consistency for consonants as assessed at the level of individual sounds and as aggregated over total sounds, three developmental sound classes, two class features, two voicing features, and five manner features. As indicated in the %C columns of the numeric section of Panel A, consonant error consistency for younger children was in the low 80% range (DAS: 83.8%, SD: 82.4%). Consonant error consistency for the older children was approximately 8-9 percentage points higher (DAS: 91.0%, SD: 91.6%). Three other sets of error consistency profiles assessing vowels, consonants-diacritics, and vowels-diacritics also did not yield significant between-group differences in error consistency. Thus, when appropriately adjusted for number of errors, the error consistency of this sample of younger and older children with suspected DAS did not differ significantly from the error consistency of children with developmental phonological disorders of unknown origin.

Summary and Conclusions

As was found in each of the other speech analyses to this point, the speech-profile analyses have not pointed to a singular difference between children with suspected DAS and children with developmental phonological disorders of unknown origin (SD). Although there were some trends associated with severity of involvement, neither the error-target analysis nor a number of error-type analyses, including analysis of error consistency, yielded a consistent pattern of significant differences.

Methodologically, two possibilities may have attenuated findings in either or both speech analysis series. First, relative to measures, perhaps the phonological measures used to define the disorder were not appropriate or sufficiently sensitive to the range of expression of the disorder. None of the analyses have addressed children's phonological comprehension, and all analyses were based on perceptual findings, rather than on acoustic or physiological data. A second possible explanation for the negative findings for speech error profiles concerns subjects. Perhaps too few of the 7 younger and/or 7 older children with suspected DAS were true positives for this childhood speech disorder. In the following analysis series we divide the group of subjects with suspected DAS into three subsets and undertake more detailed analysis of these 14 children's individual speech and prosody-voice patterns.

Is Inappropriate Stress a Diagnostic Marker for DAS?

Rationale and Procedure

The following analyses were prompted by what appeared to be the most promising findings from the group-level prosody-voice comparisons: In comparison to the average performance of children with SD, children with suspected DAS had significantly lower average scores on appropriate linguistic stress. The goal of the analyses was to assess, at the individual child level, the degree of support for inappropriate stress as a diagnostic marker for DAS.

A series of tables was assembled for visual inspection of subject-level data on each of the variables described in the previous sections. Additional information for each child was obtained from laboratory notes termed the *Transcription Commentary*. As described previously, the entire data set had been transcribed by consensus by the first and third authors prior to independent transcription of all files by the research transcriptionist. Detailed notes and commentary had been made while transcribing both the conversational speech data and responses to an articulation test (not reported in this paper). Together with comments made by the research transcriptionist and the quantitative data, these commentaries provided profiles of the speech and prosody-voice characteristics of each of the 14 children. Presentation of findings begins with the prosody-voice data.

Inappropriate Stress as a Diagnostic Marker for DAS

Table 5 includes relevant information on each child's status on stress and the other 6 suprasegmentals. As indicated in the leftmost column, the 14 children with suspected DAS are reclassified into three groups—*inappropriate stress, questionable stress*, and *appropriate stress*—based on their stress scores from the PVSP instrument. With one exception described below, reclassification of the 14 children with suspected DAS was made solely on the basis of the summary stress percentages, ordered from most to least involvement (low to high summary percentages). The 6 children with inappropriate stress had stress that is considered a *fail* on the prosody-voice screening instrument, with scores below 80% correct. Of the 8 remaining children, 3 children had *questionable* stress, with scores ranging from 80% to less than 90%, and 4 of the 5 remaining children had *appropriate* stress, with scores of 90% or higher. As shown in the Age column, the inappropriate stress group included 1 younger child and 5 older children; the questionable stress group included 2 younger children and 1 older child, and the appropriate stress group included 3 younger children and 1 older child. Rationale for excluding Child 2 from all groups will be presented below.

As indicated in Table 5, stress scores for the 6 children with inappropriate stress ranged from 4.2% of utterances with appropriate stress for Child 4 to 70.8% of utterances with appropriate stress for Child 14. In the

JSLHR, Volume 40, 286–312, April 1997

Table 5. Reclassification of the 14 children with suspected DAS based on their stress status.

Group/ Child	Age (yrs;mos)	Stress %	Stress z[a]	Stress PV Codes[b]	Phrasing %	Phrasing z	Phrasing PV Codes	Rate %	Rate z	Rate PV Codes	Loudness %	Loudness z	Loudness PV Codes	Pitch %	Pitch z	Pitch PV Codes	Laryngeal quality %	Laryngeal quality z	Laryngeal quality PV Codes	Resonance quality %	Resonance quality z	Resonance quality PV Codes
Inappropriate stress																						
4	5;10	4.2	−7.8	15	87.5	−0.4		100.0			33.3	−4.9	18	100.0	0.2		16.7	−1.5	25	91.7	−0.4	
10	8;6	29.2	−9.3	14	87.5	0.4		95.8			100.0	0.4		100.0			100.0	1.0		70.8	−0.7	
11	9;11	40.0	−7.7	15	85.0	0.2		100.0			100.0	0.4		95.0			78.9	0.5		90.0	−0.1	
8	7;1	45.0	−7.0	15	80.0	−0.2		95.0			100.0	0.4		100.0			90.0	0.8		100.0	0.5	
13	12;11	54.2	−5.6	15, 14	75.0	−0.6		95.8			100.0	0.4		100.0			91.7	0.8		0.0	−3.7	32
14	14;11	70.8	−3.2	15, 14	95.8	1.2		100.0			91.7	0.1		100.0			100.0	1.0		0.0	−3.7	32
Questionable stress																						
6	6;3	80.0	−1.2	15[c]	55.0	−3.7	2, 3, 5	95.0			95.0	0.3		100.0	0.2		70.0	0.0		0.0	−9.6	31
1	4;10	87.5	−0.6		95.8	0.5		95.8			100.0	0.7		66.7	−12.6	19	83.3	0.3		100.0	0.4	
9	7;5	87.5	−0.8		95.8	1.2		100.0			100.0	0.4		100.0			100.0	1.0		100.0	0.5	
Appropriate stress																						
3	5;9	91.3	−0.3		91.3	0.0		100.0			95.7	0.4		100.0	0.2		78.3	0.2		91.3	−0.5	
5	6;1	91.7	−0.2		100.0	0.9		95.8			95.8	0.4		100.0	0.2		83.3	0.3		75.0	−2.1	30
12	10;2	95.8	0.5		70.8	−1.0	2	70.8		11	100.0	0.4		100.0			100.0	1.0		100.0	0.5	
7	6;3	100.0	0.5		60.0	−3.2	2, 5	100.0			80.0	−0.9		100.0	0.2		100.0	0.8		0.0	−9.6	31
2	5;0	0.0	−8.1	15	79.2	−1.2	3	8.3		9	95.8	0.4		100.0	0.2		87.5	0.5		100.0	0.4	

[a] *z* scores are the standard deviation units above and below the means of the younger or older comparison groups of children with Speech Delay (SD).
[b] PV codes are provided only for scores below one standard deviation unit from the comparison group (see text).
[c] Many of this child's PV15 codes were due to misplaced stress and prolongations.

PVSP, inappropriate stress on an utterance meets criteria for one or more of the following codes: PV13: Multisyllabic Word Stress; PV14: Reduced/Equal Stress; PV15: Excessive/Equal/Misplaced Stress; or PV16: Multiple Stress Features. As indicated in the column under stress titled PV Codes in Table 5, the code most frequently associated with inappropriate stress was PV15: Excessive/Equal/Misplaced Stress. Three types of inappropriate stress are subsumed by this code—*excessive/ equal stress, misplaced stress on one or more words within a phrase, or misplaced stress on one or more words due to a block or prolongation* (criterial behaviors for each subtype are provided in Shriberg et al., 1990). Excessive/equal stress was the predominant behavior coded as inappropriate. As described below, although Child 2's score of 0% also met the criteria for inappropriate stress, her overall prosody-voice pattern met proposed *exclusionary* criteria for suspected DAS.

The middle data column under stress titled *z* provides standard deviation units as an additional index of children's status on this variable. The entries in this column are each child's *z* score for stress, calculated from the mean and standard deviation of the appropriate younger and older comparison group of children with SD. As shown in this column, stress *z* scores for the 6 children in the inappropriate stress group ranged from –3.2 to –9.3 standard deviations from the mean of their younger (Child 4) or older (Child 10, 11, 8, 13, 14) comparison groups of children with SD. With the exception of Child 6's marginal PVSP (80.0%) and *z* score (–1.2) on stress and Child 2's clearly deviant stress score discussed below, the stress *z* scores for all other children fell within one standard deviation of the mean of their comparison group of children with SD.

Evidence From the Other Suprasegmentals

Table 5 includes information for the six other suprasegmentals assessed by the PVSP procedure—Phrasing, Rate, Loudness, Pitch, Laryngeal Quality, and Resonance Quality. These data include summary percentages, *z* scores, and code-level information (only for subjects with scores below one standard deviation unit). The data were inspected to determine if there were any additional inclusionary or exclusionary criteria relative to the possibility of using stress as a diagnostic marker for DAS. As discussed below, the remaining data in Table 5 did not suggest the need for any additional considerations relative to the three-way grouping of subjects based on stress findings. Rather, they support the heterogeneity of prosody-voice profiles that have been associated with suspected DAS.

Phrasing. Beginning with the phrasing information in Table 5, 4 of the 14 children (Children 6, 12, 7, 2) had phrasing scores below one standard deviation unit from the mean of their comparison group of children with SD. As indicated in the PV Codes column for phrasing, the most common subtypes of inappropriate phrasing for these 4 children were PV2: Sound/Syllable Repetition; PV3: Word Repetition; and PV5: More Than One Word Repetition. The high percentage of utterances containing at least one sound, syllable, or word repetition in these children likely contributed to the perception of significant speech involvement, particularly the sound/syllable repetitions. However, none of the 6 children in the inappropriate stress group had phrasing scores more than one standard deviation unit below his or her respective comparison group of children with SD.

Rate. As indicated by their absence in Table 5, *z* scores could not be computed for rate because all children in both the younger and older comparison groups of children with SD scored 100% appropriate utterances on rate (i.e., no variance). Thus, Child 2's score of only 8.3% utterances with appropriate rate was considered a significant departure from comparison children, and from the rate scores of the other 14 children with suspected DAS, particularly the 6 children with inappropriate stress. The 6 children with inappropriate stress had appropriate rates on all or nearly all of their utterances, as did all other children except Child 12 whose rate was coded PV11: Fast, on over 29% of utterances. In contrast, nearly all of Child 2's utterances met criteria for code PV9: Slow Articulation/Pause Time (i.e., fewer than two syllables per second). Child 2's profile of speech and prosody-voice errors was considered to be more consistent with dysarthric speech than with apraxic speech. That is, although deviant rate has been noted in children with suspected DAS, these data suggest that it is more appropriate to view consistently slow rate as an exclusionary criterion for suspected DAS based on inappropriate stress. Additional data for this child are included below.

Loudness and pitch. Inappropriate loudness and inappropriate pitch are seldom observed in children with SD and each was observed in only 1 of the original 14 children with suspected DAS. Child 4, who had inappropriate stress, had inappropriate loudness coded PV18: Loud, on over 66% of his utterances. Child 1, who had questionable stress, had inappropriate pitch coded as PV19: Low Pitch/Glottal Fry, on over 33% of his utterances. Thus, neither inappropriate loudness nor inappropriate pitch emerged as useful for additional inclusionary or exclusionary criteria.

Laryngeal quality. Inappropriate laryngeal quality, typically coded PV24: Rough, is observed frequently in children with SD (Shriberg & Kwiatkowski, 1994). However, in the present group of children with suspected DAS, only Child 4 had a laryngeal quality score that was below one standard deviation unit from the comparison children with SD. Over 83% of this subject's utterances were coded PV25: Strained. Thus, inappropriate laryngeal quality was not deemed a relevant suprasegmental for the present purposes.

Resonance quality. A total of 6 of the 14 children with suspected DAS had inappropriate resonance quality scores on the prosody-voice measure, with 5 of these children having scores below 1 standard deviation unit from their comparison group of children with SD. As shown in Table 5, there were three different subtypes of inappropriate resonance. Child 13 and Child 14 in the inappropriate stress group were coded PV32: Nasopharyngeal. The Transcription Commentary indicated that Child 8 in this group was also considered to be borderline or questionable on every utterance for the "hot potato" percept that characterizes PV32: Nasopharyngeal. As discussed in Shriberg et al. (1990), this percept is of enhanced pharyngeal resonance, as occurs when mandible and tongue are lowered fully to cool hot food in the mouth. Child 6 and Child 7 in the questionable and appropriate stress groups, respectively, were coded PV31: Denasal, and Child 5 in the appropriate stress group was coded PV30: Nasal. Thus, there was no one resonance pattern that clearly emerged as a possible inclusionary or exclusionary behavior, although the unique resonance quality termed nasopharyngeal was considered of potential interest.

Evidence from the Speech Characteristics

Table 6 is a summary of the results of the speech analyses, essentially comparing findings for the 6 children with inappropriate stress to findings for the other 8 children with suspected DAS and the younger and older comparison children with SD. The procedures to determine *z* scores were similar to those described for the data in Table 5.

Stress and severity of involvement. As shown in the first set of data columns in Table 6, severity of speech involvement was essentially unrelated to stress status. This conclusion is based on the observation that within each of the three stress status groups there are children with relatively mild to relatively severe speech involvement as indexed by one or more of the five severity metrics. Put differently, there is no one severity metric on which all children with inappropriate stress differ from children with questionable or appropriate stress or from children in the comparison groups of children with SD.

Stress and error types. The second analysis series summarized in Table 6 assessed whether children with inappropriate stress have a different pattern of speech error types. Specifically, do they make proportionally more omission, substitution, or distortion errors than children with questionable or appropriate stress or children with SD? Error types are provided for two summary-level sets of variables: as aggregated by developmental sound class (Early-8, Middle-8, and Late-8) and Total, and as grouped by manner features (Nasal, Glide, Stop, Affricate, Fricative, Liquid). The letters *O* (Omission), *S* (Substitution), and *D* (Distortion) in each column indicate that a child's percentage of this error type was greater than one standard deviation from the appropriate comparison group of children with SD. Thus, a letter in a variable column indicates that this error type (O, S, or D) occurred proportionally more often than found in the conversational speech of children with SD.

The first error type analysis in Table 6, by developmental sound class, indicates that the error types of children with inappropriate stress do not differ from those with SD or children with questionable or appropriate stress. There is no one error type—omission, substitution, or distortion—that occurs more often for all 6 children for either Early-8, Middle-8, or Late-8 sounds. The closest finding for potential error type patterns in children with DAS is within the Early-8 sounds where 4 of the 6 children have proportionally more distortion errors than children with SD. Only one child (Child 12) of the children with questionable and appropriate stress has proportionally more distortion errors on the Early-8 sounds. This finding suggested that distortion errors on the developmentally "easier" sounds may be the locus of a speech error-type difference for children with inappropriate stress. Because the Early-8 sounds include two of the three *nasals* (/m/, /n/), the two *glides* (/j/, /w/), and three of the six *stops* (/b/, /d/, /p/), this finding was important to pursue using a manner features analysis.

The second error-type analysis in Table 6 provides information at the level of manner features, ordered from left-to-right in a generally ontogenetic sequence (cf. Shriberg, 1993). As in the analysis by developmental sound class, the letters in each cell indicate a relative error type that was more than one standard deviation above the mean of the age-appropriate comparison group of children with SD. The pattern of letters suggested that distortion of glides was associated with stress status, and hence a candidate for a diagnostic marker. All 6 of the children with inappropriate stress also had more distortions on glides than their same-aged children in the comparison groups with SD. However, 3 of the 8 children who did not have inappropriate stress also had proportionally more distortion errors on glides.

In addition to the error-type analyses shown in Table 6, we inspected individual patterns of substitutions and distortions for each of the 14 children. There were no substitution patterns or distortion types that were clearly and uniquely evident in the profiles of the 6 children with inappropriate stress. Specifically, both

Table 6. Speech characteristics of children classified as inappropriate stress, questionable stress, and appropriate stress.

Group/ Child	Severity indices[a]										Error type[c]										Error consistency	
	PCC		PCC-A		PCC-R		PVC		II		By developmental sound class				By manner feature						PCE[e]	
	%	*z*[b]	%	*z*	%	*z*	%	*z*	%	*z*	Early-8	Middle-8	Late-8	Total	Nasal	Glide	Stop	Affri-cate	Frica-tive	Liquid	%	*z*
Inappropriate stress																						
4	44.6	−2.2	46.4	−2.2	52.8	−2.0	86.7	−1.3	86.3	−0.7			O	O		D			O	O	75.0	−0.7
10	82.6	0.4	86.3	−0.2	91.0	−0.4	80.9	−2.9	98.2	0.3	D	D			D	D	D			D	100.0	0.8
11	68.8	−1.6	71.0	−3.1	72.4	−5.0	85.6	−1.7	89.8	−3.7			O	O		D		—[d]	O		71.4	−1.9
8	87.2	1.0	93.3	1.1	96.4	1.0	83.5	−2.3	97.3	−0.1	D	D	D	D		D	O	D	S	D	100.0	0.8
13	80.5	0.1	85.7	−0.3	87.9	−1.1	94.8	0.6	91.7	−2.8	D	S				D	D		S		88.9	−0.3
14	82.3	0.3	88.7	0.2	92.1	−0.1	82.0	−2.6	89.2	−3.9	D		D		D	D	D		D		—[f]	—[f]
Questionable stress																						
6	65.2	0.3	66.7	−0.3	68.7	−0.4	79.3	−3.4	79.6	−1.5				O	O	D			O		86.8	0.4
1	43.5	−2.3	43.8	−2.5	47.9	−2.5	85.6	−1.6	66.8	−3.2			O	O	S			O	O	O	86.0	0.4
9	51.0	−4.0	52.5	−6.5	55.4	−9.3	76.0	−4.1	61.1	−17.3	O		O	O		D		—[d]	O	O	85.5	−0.6
Appropriate stress																						
3	52.9	−1.2	53.7	−1.5	58.6	−1.4	75.9	−4.3	40.2	−6.5			O	O				O	O	O	83.3	0.1
5	62.7	0.0	66.5	−0.3	72.2	0.0	83.5	−2.2	84.4	−0.9		S				S	S	D			71.4	−1.1
12	89.4	1.3	90.8	0.6	92.5	0.0	96.0	0.9	98.4	0.5	D		S			D	D	—[d]	S	S	100.0	0.8
7	76.1	1.6	77.0	0.8	79.4	0.7	81.3	−2.8	95.1	0.4		S				S					100.0	1.7
2	51.9	−1.3	55.8	−1.3	58.1	−1.5	82.4	−2.5	88.9	−0.3	S		O				S		O		84.0	0.2

[a]PCC: Percentage Consonants Correct; PCC-A: Percentage Consonants Correct-Adjusted; PCC-R: Percentage Consonants Correct-Revised; PVC: Percentage Vowels Correct; II: Intelligibility Index
[b]*z* scores are the standard deviation units above and below the means of the younger or older comparison groups of children with Speech Delay (SD).
[c]The letters *O* (Omission), *S* (Substitution), and *D* (Distortion) indicate an error type percentage that is more than one standard deviation above the mean of that error type by children in the younger or older comparison group with SD.
[d]Omission errors were present but, because the older comparison group with SD had no omission errors, there was no standard deviation.
[e]PCE: Percentage Consistent Errors
[f]Unable to calculate due to speech sample characteristics.

within-class substitutions (e.g., stops for stops) and out-of-class substitutions (e.g., fricatives for stops) occurred for all three subgroups within the 14 children with suspected DAS. Moreover, as divided into diacritics reflecting changes (i.e., distortions) of *manner, voicing, place, force*, and *duration*, there were no quantitative or qualitative differences in the error types of the 6 children with inappropriate stress in comparison to children with questionable or appropriate stress. It should be noted that these analyses used conservative decision criteria.

The final error-type analyses shown in Table 6 are the Percentage of Consistent Errors (PCE) scores and z scores for children in the three stress status groups. As shown, PCE was not associated with stress status. Of the 2 children scoring below one standard deviation from their comparison group of children with SD, one (Child 11) was in the inappropriate stress group and the other (Child 5) in the appropriate stress group.

Summary

The goal of this study was to attempt to identify, from analysis of conversational speech samples, at least one characteristic of children with suspected DAS that differentiated them from children with speech delay. What was sought was one or more behaviors that had construct validity relative to the literature on DAS, and that also could meet tests of divergent criterion validity relative to the occurrence of the behavior(s) in appropriate comparison groups.

The initial analysis of the 14 children with suspected DAS using grouped means and standard deviations suggested that inappropriate stress was the only characteristic that might meet both goals. Inappropriate stress meets construct validity criteria as being wholly consistent with the clinical percept of DAS as a term used for children who "sound different." Divergent criterion validity support is provided from the many statistical comparisons of children with DAS to children with SD, with inappropriate stress the only characteristic significantly differentiating the two groups.

These data are interpreted as support for a proposal that inappropriate stress is a diagnostic marker for a subtype of DAS. Accordingly, only 6 of the 14 children (43%) with suspected DAS were viewed as true positives for this subtype, with 3 remaining children (21%) considered questionable and the other 5 (36%) considered to not meet this marker criterion. The proposal that inappropriate stress is a diagnostic marker for a subtype of DAS has attractive implications for theory and practice. The neurodevelopmental correlates of linguistic stress provide fertile ground for explanatory accounts and, from an applied perspective, perceptual and instrumental measures for sensitive and reliable assessment of stress can readily be developed. However, before considering such issues, the external validity of the findings of Study I were assessed in two cross-validation studies. Findings from an internal validation study are reported next, with findings for an external cross-validation study reported in a companion paper (Shriberg et al., 1997b).

Study II: Internal Cross-Validation of Stress as a Diagnostic Marker for DAS

Study II is a retrospective cross-validation study using a sample of children referred to a phonology clinic for suspected DAS. The primary goal was to assess the level of support for inappropriate stress as a diagnostic marker of DAS. Secondary goals were to follow up on several other prosody-voice findings and trends from Study I.

Method

Table 7 provides descriptive information for 20 children with suspected DAS. The 20 children include 14 children referred to the University of Wisconsin Phonology Clinic from 1989–1993 (referenced previously in Shriberg et al., 1997a, Table 1), plus 6 additional children with suspected DAS referred to the clinic prior to 1989. The 15 boys and 5 girls ranged in age from 3 years 2 months to 9 years 3 months (M = 6 years 1 month) at the time the prosody-voice data in Table 7 were gathered. Using 7 years as the dividing point, as in Study I, the children were divided into 12 younger children with suspected DAS and 8 older children with suspected DAS. As shown in Table 7, at the time of the assessment or intervention report the 20 children had received from approximately 7 months to as much as 6 years 3 months of speech-language treatment elsewhere or at the Phonology Clinic.

Conversational speech samples meeting research criteria were not available for phonetic transcription or prosody-voice coding for most of these 20 clinic referrals. However, prosody-voice information was available in the diagnostic and intervention reports completed by author JK, who completed speech-language assessments for all 20 children. The extensive diagnostic and intervention reports included summaries of each child's speech status, as well as status on each of the seven prosody-voice suprasegmentals. Although methodology for the Prosody-Voice Screening Profile (PVSP) procedure was not completed until 1990, prior versions of the procedure using generally similar response definitions

Table 7. Descriptive and prosody-voice data for 20 children with possible or suspected DAS referred to a university phonology clinic in Madison, Wisconsin.

				Prosody-voice status[a]						
				Prosody			Voice			
Child	Age at assessment (yrs;mos)	Gender	Approximate amount of treatment prior to speech sample (yrs;mos)	Phrasing	Rate	Stress	Loudness	Pitch	Laryngeal quality	Resonance quality
Younger subgroup										
1	3;2	M	1;0	A	A	A	A	A	A	A
2	3;2	M	0;7	CNJ	Q (fast)	I	A	A	A	A
3	4;3	M	1;1	A	A	A	A	A	A	A
4	4;7[b]	M	2;10	A	I (slow)	I	A	A	A	A
5	4;10	M	1;8	I	A	A	A	A	A	A
6[c]	4;11	F	1;11	A	A	I	Q	A	I	A
7	5;4	F	2;4	A	I (slow)	I	A	A	A	A
8[c]	5;8	F	1;3	I	A	I	A	A	I	A
9	5;8	M	1;5	A	Q (fast)	I	A	A	A	A
10	6;6	M	3;6	A	A	I	A	A	A	I
11	6;7	M	4;7	I	A	I	A	A	A	A
12	6;11	M	4;11	A	A	A	A	A	A	A
Older subgroup										
13	7;0	M	4;0	Q	Q (fast)	Q	A	A	A	A
14	7;1	M	4;1	A	I (fast)	I	A	A	A	Q
15	7;1	M	4;0	A	Q (fast)	CNJ	A	A	A	A
16	7;2	F	1;6	I	A	A	A	A	A	A
17[c]	7;3	M	3;5	Q	A	I	A	A	I	A
18[c]	7;5	M	4;5	A	Q (fast)	A	Q	A	I	A
19[c]	7;9	M	3;2	Q	A	I	I	A	I	A
20[c]	9;3	F	6;3	A	I (slow)	Q	Q	A	A	A
M	6;1	%Male: 75	%A:	63	55	32	80	100	75	90
SD	1;7	%Female: 25	%Q:	16	25	11	15	0	0	5
			%I:	21	20	58	5	0	25	5

[a]A: Appropriate; Q: Questionable; I: Inappropriate; CNJ: Could Not Judge (due to telegraphic speech or intelligibility deficits). Inappropriate Rate entries are indicated as too slow or too fast.
[b]Also referred for possible dysfluency.
[c]Referred prior to 1989.

to those in Shriberg et al. (1990) had been used for all diagnostic and intervention reports (cf. Shriberg & Kwiatkowski, 1982; Shriberg & Widder, 1990; Shriberg et al., 1986). Based on these case records data, children's status on each suprasegmental was trichotomized in the clinical reports as *appropriate, questionable*, or *inappropriate*. Intrajudge and interjudge prosody-voice coding reliability was established during this period using conversational speech samples from 28 3- to 19-year-old children selected from a database of children with normally developing and delayed speech. At the summative (i.e., pass/fail) level of judgment, intrajudge and interjudge agreement (with the first author and the transcriber in Study I) on the seven prosody-voice variables ranged from point-to-point percentages in the low 70s to 100% agreement (Shriberg et al., 1992).

Results

The percentages at the bottom of Table 7 summarize the proportion of younger and older children with suspected DAS coded Appropriate (*A*), Questionable (*Q*), or Inappropriate (*I*) on each of the prosody-voice variables. It is efficient to divide the findings for the seven suprasegmentals in Table 7 into three sets of observations, beginning with findings for a set of five variables.

Phrasing, Loudness, Pitch, Laryngeal Quality, and Resonance Quality

The percentage distributions for Phrasing, Loudness, Pitch, Laryngeal Quality, and Resonance Quality

JSLHR, Volume 40, 286–312, April 1997

as shown in Table 7 are similar to those found in Study I (see Figure 1) and to findings for children with normal and delayed speech (Shriberg & Kwiatkowski, 1994, Table 3). Beginning with the phrasing findings from the latter reference, 52% of the 71 3- to 5-year-old speech-normal children had appropriate phrasing and 71% of the 62 3- to 6-year-old speech-delayed children had appropriate phrasing. In the present study, 63% of the 20 children with suspected DAS had appropriate phrasing, which falls nearly in between the values for the normal and speech-delayed children in Shriberg and Kwiatkowski (1994). As described in the technical references for the PVSP (Shriberg et al., 1990, 1992), procedures to code phrasing are sensitive to language processing variables, and are similar to those used to code mazing (cf. Miller & Chapman, 1986). Thus, questionable or inappropriate phrasing does not distinguish children with suspected DAS from children with normal speech or from children with speech delay of unknown origin.

The percentages of children with appropriate loudness (80%), pitch (100%), laryngeal quality (75%), and resonance quality (90%) in Table 7 are also comparable to the percentages for normal-speaking and speech-delayed children in Shriberg and Kwiatkowski (1994, Table 3). It should be noted that the ages of the children in the two groups in Shriberg and Kwiatkowski are comparable only to the ages of children in the younger group in Table 7. However, note that the percentages of children with questionable and inappropriate entries for the five prosody-voice variables are distributed approximately equally across the younger and older age groups. Thus, the data in Table 7 suggest that these five prosody-voice variables are also not useful diagnostic markers of DAS. Notably, the Study I trend for some children with suspected DAS to have nasopharyngeal resonance was not supported in the present data set.

Rate

A second observation regarding the entries in Table 7 concerns the data for rate. Only 55% of children met criteria for appropriate rate, with 25% having questionable rate and 20% having inappropriate rate. Recall that, in Study I, 12 of the 14 children (86%) scored above 80% for rate on the PVSP. The only child in Study I with significantly slow rate (Child 2) was considered to have a form of dysarthric speech. Differences in the percentage of children with questionable and inappropriate rate in Study II versus Study I may be associated with the more well-developed PVSP procedures in Study I, which require stop-watch verification of syllable rates. Whatever the reason, it is important to observe the direction of questionable or inappropriate rate, as indicated in parentheses in Table 7. Of the 9 children who were coded as questionable or inappropriate rate in Table 7, 3 were too slow and 6 were too fast. Rate differences were distributed approximately equally across the younger and older groups, with instances of each rate difference type in each age group. If rate is eventually implicated as a marker for DAS, the underlying pathogenic processes would have to account for speech rates that are alternatively too slow and too fast.

Stress

The crucial statistics in Table 7 for the present purposes are the percentages of children with suspected DAS with inappropriate stress in conversational speech. As shown in Table 7, inappropriate stress was coded for 11 (58%) of the 19 children whose stress could be judged and questionable stress for an additional 2 children (11%). In comparison to these findings, only 1.4% of 3- to 5-year-old normally speaking children and 17.8% of 3- to 6-year-old children with speech delays of unknown origin (Shriberg & Kwiatkowski, 1994, Table 3) had either inappropriate or questionable stress.

Two associated considerations increase the significance of the stress findings for Study II. First, as shown in Table 7, 5 of the 13 children with inappropriate or questionable stress were in the older subgroup of children with suspected DAS. As discussed in the third paper in this series (Shriberg et al., 1997b), one perspective on inappropriate stress is that it should normalize with advancing age. Second, the 5 children in this older subgroup had received many years of treatment by the time of these reports, ranging from 3 years 2 months to 6 years 3 months. Both advanced age and more extensive intervention experience militate against finding persisting stress involvement.

Discussion

Before considering some information about linguistic stress that may bear on the present findings, it is appropriate to address several measurement issues.

Measurement Issues

Sensitivity of Speech and Prosody-Voice Measures

Both the nonsignificant speech findings in Study I and the significant stress findings in both Study I and Study II could be associated with sensitivity of measurement issues at all stages of conversational sampling. More fine-grained procedures to sample and quantify children's knowledge or productive control of English stress might have yielded stress differences for some to all of the children with suspected DAS in the three studies. Moreover, alternative ways to sample consistency of speech errors might have identified more inconsistency or variability in children with suspected DAS. Until the present findings are cross-validated using other measurement approaches, the possibility of sensitivity constraints or Type II measurement errors represent a potential threat to the internal validity of findings. However, several considerations support the adequacy of the sampling and data reduction procedures used in the present studies.

First, conversational speech sampling has optimum face validity for the questions asked about speech errors and phrasal stress (see later discussions of stress). Although minimally directed conversation may not fully press children's speech and prosody-voice capability, appropriately obtained conversational samples have been documented to be sensitive to the presence of clinically significant deficits in verbal communication (Morrison & Shriberg, 1992; Shriberg & Kwiatkowski, 1985).

Second, although physiologic or acoustic data may offer increased sensitivity to articulatory or spectral characteristics of speech and prosody-voice, the narrow phonetic transcription and prosody-voice coding procedures used in the present study were developed expressly for fine-grained analysis of children's speech. As discussed in the third paper in this series, more sensitive measurement procedures using controlled linguistic stimuli, acoustically aided transcription procedures, and alternative phonological analyses might maximize the sensitivity of assessment procedures to inconsistent speech production and inappropriate stress. However, given the individual and total sample sizes in these two studies, the measurement procedures used in the present studies are viewed as having sufficient sensitivity to support the descriptive-explanatory discussions to follow.

Specificity of Speech and Prosody-Voice Measures

A second measurement issue is whether the inappropriate stress findings for some children with suspected DAS are veridical reflections of only the stress domain. If the procedure used to assess stress lacks specificity, stress deficits may actually reflect involvements in other domains, with consequent implications for interpretation of findings. Specificity issues concern the response definitions used for utterances coded as PV15: Excessive/Equal Stress, which was the most frequent source of inappropriate stress in the three studies. Support for the specificity of the stress findings is based on the following two considerations.

First, the possibility of confounds between phonologic-morphosyntactic levels and prosodic-vocal levels was directly addressed at each stage of the development of the prosody-voice screening instrument (Shriberg et al., 1990, 1992). For the stress domain, it was anticipated that coding criteria needed to be developed to differentiate utterances coded PV15 from utterances in which stress was appropriate. For example, children with frequent final consonant deletions, numerous disfluencies, erratic phrasing, or morphosyntactic deficits yielding "telegraphic" speech might be falsely perceived as having excessive and/or equal stress. To prevent such confounds in assessing segmental and suprasegmental levels of phonology, a table of instructions requires the prosody-voice coder to attend to the frequency, duration, and intensity of vowels, and to observe coarticulatory phenomena at word boundaries. Training exemplars and other instructional examples in the audio-training series for this instrument contrast utterances from children with excessive/equal stress with utterances from children with impoverished phonetic inventories, frequent final consonant deletion, frequent repetitions, revisions, and other phrasing irregularities, but normal phrasal stress.

A second consideration supporting the specificity of the stress findings is that PV15: Excessive/Equal Stress was observed in children who, on inspection, had fully complete phonetic inventories, age-appropriate syntax, and low percentages of final consonant deletion. For example, the stress differences noted in the children with suspected DAS in the older groups included children whose PCC scores were in the high 80s. Thus, (a) stress was measured with procedures that ensured orthogonality between stress deficits and all other segmental and suprasegmental domains, and (b) stress deficits were observed in children at all levels of severity of speech involvement.

From these perspectives, the stress data in the present study are considered to have been obtained by measures with adequate sensitivity and specificity. Additional comment on measurement issues is deferred to the final paper in this series.

Inappropriate Linguistic Stress

Summary of Findings

The present studies used a three-domain system of stress marking to assess children's stress in conversational speech: (a) *lexical stress*—stressing and destressing appropriate syllables in a word, (b) *phrasal* (or *sentential*) *stress*—stressing and destressing syllables and words according to their morphological and syntactic function in a phrase, and (c) *emphatic stress*—stressing and destressing syllables and words for meaning and emotive functions. Based on utterance-by-utterance perceptual coding, inappropriate *phrasal* stress was found in 43% (6/14) of the children with suspected DAS in

Study I, and 58% (11/19) of the children with suspected DAS in Study II. In comparison, inappropriate stress was observed in only 10% of two large groups of children with normal speech or speech delay. For some of the children with suspected DAS, instances of inappropriate *lexical* stress were also informally observed on articulation test responses not included in this paper.

The crucial observation about the stress findings is that the type of inappropriate phrasal stress observed most often met criteria for Prosody-Voice Code 15: Excessive/Equal Stress, as defined in the perceptual system to assess prosody-voice (Shriberg et al., 1990). Although the response definitions for PV Code 15 do not require the excessive, exaggerated prosody termed *hyperprosody* by Monrad-Krohn (1963), the criteria for the code are entirely consistent with the three "dysprosodic patterns" noted by Kent and Rosenbek (1983) to characterize adult AOS: articulatory prolongations, syllable segregation, and lengthening of normally unstressed vowels.

More generally, inappropriate stress of the two types found in the present studies—inappropriate lexical stress and inappropriate phrasal stress—are widely attested findings in studies of adults with AOS. For example, Odell, McNeil, Rosenbek, and Hunter (1990) found that syllabic stress errors at the single-word production level were characteristic of their adult subjects with AOS, but not subjects with conduction aphasia. Square-Storer, Darley, and Sommers (1988) characterize adults with AOS as having monotonous, effortful speech with equal stress and poor volume control. Kent and Rosenbek (1983) noted that a key prosodic feature in their adult subjects was failure to destress unstressed syllables and function words such as "the," "in," "on," and "was." They noted that these speakers produced a flattened intensity envelope across sequences of syllables that normally would be both stressed and unstressed. Wertz, LaPointe, and Rosenbek (1984) note that adults with AOS "sound like they give each syllable in their utterances equal stress" (p. 70). McNeil and Kent (1990), summarizing findings in adult AOS (and Broca aphasia), note that "…syllable relief tends toward temporal regularity and amplitude uniformity….These acoustic features should correlate with neutralization of stress pattern and dysrhythmia, that is, with dysprosody" (p. 364).

Stress deficits have also been amply documented as evident in children with suspected DAS from their earliest to latest periods of speech acquisition. Tate's (1991) description of early stress deficits in three infants later considered to have DAS is instructional. All three children were "quiet" babies whose vocalizations were reportedly limited in character and length. Early vocalizations appeared not to be "shaped" by adults around them. Rather, the infants had "flat" intonations that did not improve with increases in phonetic inventory. The classic study of older children with suspected DAS by Yoss and Darley (1974) reported that these children's spontaneous speech had a "measured effect…with a tendency toward equalization of stress" (p. 412). Robin, Hall, Jordan, and Gordan (1991) found that the stress differences of children with suspected DAS included changes in all three parameters of the speech wave (i.e., frequency, amplitude, and duration). Colson (1988) found reduced stressed-unstressed vowel duration ratios for three children with suspected DAS. These and many other studies suggest that, as with adult AOS, children with suspected DAS have primary difficulty destressing unstressed syllables, yielding a pattern characterized as lacking prosodic contrast. However, this pattern does not always include the vowel prolongations, excessive loudness, and long pauses reported for adults with AOS, although Shuster, Ruscello, and Haines (1989) report "articulatory prolongation, syllable segregation, and longer segment durations" (p. 10) in a 15-year-old child with suspected DAS studied acoustically.

Summary and Conclusions

The findings for Study I and Study II are viewed as supporting a conclusion that inappropriate stress may be a diagnostic marker for developmental apraxia of speech. Using inappropriate stress as a candidate diagnostic marker, the true positive rate for the 20 children referred for suspected DAS in Study II would be 58%, but the false positive rate would be approximately 42% (with rounding differences). These figures are within 15 percentage points of those found in Study I, which found approximately 43% true positives. Due to the relatively small sample sizes in Studies I and II, the 15-point difference in true positive rates could reflect differences in classification assignments for as few as 2–3 children.

Several design limitations diminish the value of Study II as a cross-validation of Study I. From a measurement perspective, evaluation of prosody-voice was completed by a clinician-researcher who in all cases was aware of the suspicion of DAS, and who completed the evaluation using a prior version of the prosody-voice procedure. From a demographic perspective, the 20 children referred for a concern about or suspected DAS were all obtained from the same geographic area, reflecting the referral concerns of clinicians in one midwestern community. And most importantly, in Study II there were no speech data available for close examination. The following paper reports a study designed to address these internal and external validity concerns, followed by a consideration of theoretical, research, and clinical issues.

Acknowledgments

This research was supported by a grant from the U.S. Public Health Service, NIDCD DC00496. We express sincere thanks to the following people for their time and capable research assistance at different stages of this project: Barrie Babow, Mary Coberly, Barbara Ekelman, Frederic Gruber, Kit Hoffmann, Patricia Hovel, Maureen McGowan Jepsen, Gregory Lof, Jane McSweeny, Amparo Ortez, Dorothy Ross, Carmen Rasmussen, Carol Widder, and David Wilson. Special thanks to Diane Austin for her insightful assistance with all phases of the computational analyses and manuscript processing.

References

Aram, D. M., & Horwitz, S. J. (1983). Sequential and non-speech praxic abilities in developmental verbal apraxia. *Developmental Medicine and Child Neurology, 25*, 197–206.

Arthur, G. (1952). *The Arthur Adaptation of the Leiter International Performance Scale*. Washington, DC: Psychological Service Center Press.

Carterette, E., & Jones, M. (1974). *Informal speech: Alphabetic and phonemic texts with statistical analyses and tables*. Berkeley, CA: University of California Press.

Colson, K. A. (1988). *Perception and production of changes in stress placement, pitch and duration by patients with focal left cortical insult*. Unpublished doctoral dissertation, University of Iowa, Iowa City.

DiSimoni, F. (1978). *The Token Test for Children*. Boston, MA: Teaching Resources.

Dunn, L. M. (1965). *Expanded manual for the Peabody Picture Vocabulary Test* (Rev. ed.). Circle Pines, MN: American Guidance Service.

Elston, R. C., & Namboodiri, K. K. (1980). Types of disease and models for their genetic analysis. *Schizophrenia Bulletin, 6*, 368–374.

Foster, R., Giddian, J. J., & Stark, J. (1973). *Manual for the assessment of children's language comprehension*. Palo Alto, CA: Consulting Psychologists Press.

Gardner, M. F. (1979; 1983). *Expressive One-Word Picture Vocabulary Test: Upper Extension*. Novato, CA: Academic Therapy Publications.

Hoffmann, K. (1982). *Speech sound acquisition and natural process occurrence in the continuous speech of three- to six-year-old children*. Unpublished master's thesis, University of Wisconsin-Madison.

Irwin, J. V., & Wong, S. P. (1983). *Phonological development in children 18 to 72 months*. Carbondale, IL: Southern Illinois University Press.

Kent, R. D., & Rosenbek, J. C. (1983). Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research, 26*, 231–249.

Kwiatkowski, J., & Shriberg, L. D. (1993). Speech normalization in developmental phonological disorders: A retrospective study of capability-focus theory. *Language, Speech, and Hearing Services in Schools, 24*, 10–18.

Lee, L. L. (1964). *Northwestern Syntax Screening Test*. Evanston, IL: Northwestern University Press.

Mader, J. B. (1954). The relative frequency of English consonant sounds in words in the speech of children in grades one, two and three. *Speech Monographs, 21*, 249–300.

McNeil, M. R., & Kent, R. D. (1990). Motoric characteristics of adult aphasic and apraxic speakers. In G. E. Hammond (Ed.), *Cerebral control of speech and limb movements* (pp. 349–386). New York: Elsevier.

Miller, J. F., & Chapman, R. (1986). *Systematic analysis of language transcripts (SALT)*. University of Wisconsin-Madison.

Mines, M., Hanson, B., & Shoup, J. (1978). Frequency of occurrence of phonemes in conversational English. *Language and Speech, 21*, 221–241.

Monrad-Krohn, G. H. (1963). The third element of speech: Prosody and its disorders. In L. Halpern (Ed.), *Problems in dynamic neurology* (pp. 101–118). Jerusalem: Hebrew University Press.

Morrison, J. A., & Shriberg, L. D. (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*, 259–273.

Odell, K., McNeil, M. R., Rosenbek, J. C., & Hunter, L. (1990). Perceptual characteristics of consonant productions by apraxic speakers. *Journal of Speech and Hearing Disorders, 55*, 345–359.

Pennington, B. F. (1986). Issues in the diagnosis and phenotype analysis of dyslexia: Implications for family studies. In S. D. Smith (Ed.), *Genetics and learning disabilities* (pp. 69–96). San Diego, CA: College-Hill Press.

Robin, D. A., Hall, P. K., Jordan, L. S., & Gordan, A. J. (1991, November). *Developmentally apraxic speakers' stress production: Perceptual and acoustic analyses*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Atlanta, GA.

Shriberg, L. D. (1983). The method of natural phonological processes. In W. Perkins (Ed.), *Current therapy of communication disorders* (pp. 3–9). New York: Thieme-Stratton.

Shriberg, L. D. (1986). *PEPPER: Programs to examine phonetic and phonologic evaluation records*. Hillsdale, NJ: Lawrence Erlbaum.

Shriberg, L. D. (1991). Directions for research in developmental phonological disorders. In J. F. Miller (Ed.), *Research on child language disorders: A decade of progress* (pp. 267–276). Austin, TX: Pro-Ed.

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., Aram, D. M., & Kwiatkowski, J. (1997a). Developmental apraxia of speech: I. Descriptive perspectives. *Journal of Speech, Language, and Hearing Research, 40*, 273–285.

Shriberg, L. D., Aram, D. M., & Kwiatkowski, J. (1997b). Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language, and Hearing Research, 40*, 313–337.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (in press-a). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (in press-b). The Speech Disorders Classification System (SDCS): Extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research*.

Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J. (1994). Developmental phonological disorders III: Long-term speech-sound normalization. *Journal of Speech and Hearing Research, 37*, 1151–1177.

Shriberg, L. D., & Kent, R. D. (1995). *Clinical phonetics* (2nd ed.). Boston: Allyn & Bacon.

Shriberg, L. D., & Kwiatkowski, J. (1980). *Natural process analysis: A procedure for phonological analysis of continuous speech samples*. New York: Macmillan.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders I: A diagnostic classification system. *Journal of Speech and Hearing Disorders, 47*, 226–241.

Shriberg, L. D., & Kwiatkowski, J. (1983). Computer-assisted natural process analysis (NPA): Recent issues and data. *Seminars in Speech, Language, and Hearing, 4*, 389–406.

Shriberg, L. D., & Kwiatkowski, J. (1985). Continuous speech sampling for phonologic analyses of speech-delayed children. *Journal of Speech and Hearing Disorders, 50*, 323–334.

Shriberg, L. D., & Kwiatkowski, J. (1994). Developmental phonological disorders I: A clinical profile. *Journal of Speech and Hearing Research, 37*, 1100–1126.

Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B. (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51*, 140–161.

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

Shriberg, L. D., Kwiatkowski, J., & Rasmussen, C. (1990). *The Prosody-Voice Screening Profile*. Tucson, AZ: Communication Skill Builders.

Shriberg, L. D., Kwiatkowski, J., Rasmussen, C., Lof, G. L., & Miller, J. F. (1992). *The Prosody-Voice Screening Profile (PVSP): Psychometric data and reference information for children* (Tech. Rep. No. 1). Madison, WI: Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison.

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics and Phonetics, 5*, 225–279.

Shriberg, L. D., & Widder, C. J. (1990). Speech and prosody characteristics of adults with mental retardation. *Journal of Speech and Hearing Research, 33*, 627–653.

Shuster, L. I., Ruscello, D. M., & Haines, K. B. (1989, November). *Acoustic patterns of developmental apraxia of speech*. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, St. Louis, MO.

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

Spriestersbach, D. C., Morris, H. L., & Darley, F. L. (1978). Speech mechanism examination. In F. L. Darley & D. C. Spriestersbach, *Diagnostic methods in speech pathology* (pp. 322–345). New York: Harper & Row.

Square-Storer, P., Darley, F. L., & Sommers, R. (1988). Nonspeech and speech processing skills in patients with aphasia and apraxia of speech. *Brain and Language, 33*, 65–85.

Stoel-Gammon, C., & Dunn, C. (1985). *Normal and disordered phonology in children*. Baltimore: University Park Press.

Tate, M. (1991, August). *A longitudinal study of the speech acquisition of three siblings diagnosed as verbally dyspraxic*. Paper presented at the XIIth International Congress of Phonetic Sciences, Aix-en-Provence, France.

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

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children* (Rev.). New York: The Psychological Corporation.

Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management*. New York: Grune & Stratton.

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.

Yoss, K. A., & Darley, F. L. (1974). Developmental apraxia of speech in children with defective articulation. *Journal of Speech and Hearing Research, 17*, 399–416.

Received June 4, 1996

Accepted October 13, 1996

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