CLINICAL LINGUISTICS & PHONETICS, 2003, VOL. 17, NO. 7, 575–595 Taylor & Francis healthsciences

A diagnostic marker for childhood apraxia of speech: the coefficient of variation ratio

LAWRENCE D. SHRIBERG †, JORDAN R. GREEN †‡, THOMAS F. CAMPBELL §, JANE L. MCSWEENY † and ALISON R. SCHEER †

† Waisman Center, University of Wisconsin-Madison, WI, USA
‡ Department of Communicative Disorders, University of Wisconsin-Madison, WI, USA
§ Children's Hospital of Pittsburgh, Pittsburgh, PA, USA

(*Received 12 August 2002; accepted 30 October 2002*)

Abstract

Terms such as *isochrony*, *syllable segregation*, *scanning speech* and *staccato-like rhythmic quality* have been used to characterize the temporal regularity that may be a core feature of apraxia of speech. The present report describes a procedure to quantify temporal regularity in children with suspected apraxia of speech (sAOS). Conversational speech samples from 15 such children, together with samples from 30 3–6-year-old children with normal speech acquisition and 30 3–6-year-old children with moderate to severe speech delay of unknown origin, were selected from an audio archive. Signal processing routines were developed to identify and measure the duration of speech and pause events in 24 utterances from the speech samples of each of the 75 speakers. A value termed the *coefficient of variation* expressed the normalized variability in the durations of each participant's speech events and pause events within each utterance. A metric termed the *coefficient of variation ratio*, derived by dividing the coefficient of variation for pause events by the coefficient of variation for speech events, expressed a speaker's relative temporal variation in the two domains. The 15 children with sAOS had higher coefficient of variation ratios than the 30 children in each of the two comparison groups, indicating that the children with sAOS had proportionally more variation in the duration of pause events and/or less variation in the duration of speech events. Findings are interpreted as supporting the view that a constraint in speech timing is a core feature of the praxis disorder that defines a developmental form of apraxia of speech.

Keywords: Apraxia, assessment, development, speech disorders, speech timing.

Introduction

Childhood apraxia of speech (CAS) is a rare disorder that takes its name from the presumption of a deficit in speech praxis. A companion paper reviewed issues and

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

Clinical Linguistics & Phonetics
ISSN 0269-9206 print/ISSN 1464-5076 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0269920031000138141

empirical findings for a diagnostic marker of stress that were viewed as consistent with the core deficit of speech praxis (Shriberg, Campbell, Karlsson, Brown, McSweeny and Nadler, 2003). The present paper focuses on another descriptive feature of CAS that is also presumed to reflect the core praxis disorder—a reduction in temporal variation associated with the percept of isochronous (similarly timed) and segregated (non-coarticulated) syllables and words (cf. Shuster, Ruscello and Haines, 1989; Hall, Jordan and Robin, 1993; Shriberg, Aram and Kwiatkowski, 1997a). Related constructs in the literature on acquired apraxia in adults include *scanning speech*, *robotic speech*, *staccato-like speech* and more generally, *abnormal speech timing* (e.g. Kent and Rosenbek, 1983; Ackermann and Hertrich, 1994; Whiteside and Varley, 1998; Kent, Weismer, Kent, Vorperian and Duffy, 1999).

Some research findings from a British family

The lack of validated inclusionary criteria for childhood apraxia of speech continues to be the primary constraint on the development of a coherent descriptive-explanatory account of this proposed disorder. As suggested in the companion paper (Shriberg *et al.*, 2003), all findings in the experimental literature have to be considered tentative due to the lack of a gold standard for CAS and to the large differences in the inclusionary criteria used by investigator groups. The majority of empirical studies and evaluative reviews view CAS as a movement disorder, but other perspectives include the possibility of significant, if not core, deficits at 'higher' levels of linguistic representation and speech processing (cf. Odell and Shriberg, 2001). Additional discussion of this ongoing dialogue is beyond the scope of the present study. A representative sample of these alternative views, which have continued during the past decade, is available in the Hall-Robin dialogue (Hall, 1992; Robin, 1992).

Landmark studies of a multigenerational British family (known as the 'KE family') have provided the first opportunity to address the origin and nature of CAS without the limitations of case-control designs referenced in the companion paper (Vargha-Khadem, Watkins, Alcock, Fletcher and Passingham, 1995; Fisher, Vargha-Khadem, Watkins, Monaco and Pembrey, 1998). Half of the members of this family have an orofacial (i.e. non-speech) apraxia that is transmitted as an autosomal dominant trait. Because the orofacial apraxia appears to cosegregate with suspected apraxia of speech, investigators have been able to address both distal and proximal explanatory processes for both disorders in affected family members. The gene transmitted in this family, FOXP2 (a transcription factor), was identified by positional cloning (Lai, Fisher, Hurst, Levy, Hodgson, Fox, Jeremiah, Povey, Jamison, Green, Vargha-Khadem and Monaco, 2000; Lai, Fisher, Hurst, Vargha-Khadem and Monaco, 2001). MRI, PET and other neuroimaging studies of affected family members have identified a number of anatomical abnormalities in cortical and subcortical motor-related areas of the frontal lobe (Vargha-Khadem, Watkins, Price, Ashburner, Alcock, Connelly, Frackowiak, Friston, Pembrey, Mishkin, Gadian and Passingham, 1998; Watkins, Gadian and Vargha-Khadem, 1999; Watkins, Dronkers and Vargha-Khadem, 2002; Belton, Salmond, Watkins, Vargha-Khadem and Gadian, 2003). It should be noted that other investigator groups have reported additional neuroanatomical sites associated with acquired apraxia of speech (see review in Kent, 2000).

Findings from studies of the orofacial praxis deficit and the apraxia of speech in

the affected members of this British family are of particular interest in the present context. Alcock, Passingham, Watkins and Vargha-Khadem (2000a; b) have reported that these individuals were not deficient in the perception or production of pitch, but that they were markedly impaired in the perception and production of rhythm, as assessed in both vocal and manual (a tapping test) modalities. These authors conclude, '… intonation abilities are not impaired in the affected family members, whereas their timing abilities are impaired. Neither their linguistic nor oral praxic deficits can be at the root of their impairment in timing; rather, the reverse may be true' (Alcock *et al.*, 2000b, p. 34). Thus, these investigators speculate that the praxis deficit that defines apraxia of speech may include at its core an impairment in speech timing.

Other than the investigators' anecdotal comments on the speech disorder in family members with orofacial apraxia, there are few available technical data on the specific patterns of speech and prosody errors of affected family members. Hurst, Baraitser, Auger, Graham and Norell (1990) provided case histories for six affected family members, focusing primarily on these speakers' difficulty sequencing movements, including impairment in the organization of speech movements and manual movements for signing. Fee (1995) provided a comprehensive description of the consonant errors of eight affected family members sampled at two points in time but did not comment on whether any of the speech errors were consistent with apraxia of speech. Goad (1998) provided a thorough analysis of a pronounced impairment in plurals in five affected adult family members, focusing on alternative (prosodic versus morphological) theoretical explanations also without comment on speech-prosody errors in relation to apraxia of speech.

A fourth paper by Piggott and Kessler Robb (1999) provided a well developed linguistic description of the prosodic characteristics of four affected family members ranging from 14–78 years in age and four approximately age-matched non-affected family members. These participants responded to a series of preference judgement and production tasks that included both real and novel multisyllabic words that assessed participants' knowledge of lexical stress rules. Piggott and Kessler Robb's findings provide the best available information on a prosodic constraint shared by affected family members. The investigators report that these family members had considerably more incorrect and variable judgements of what constituted appropriate lexical stress, even on the tasks that contained only real words. Moreover, affected family members' productions of words more than two syllables long were described as aberrant in timing. The investigators, who were well trained in phonetic transcription, noted that polysyllabic words were produced with 'prominent pauses separating them' and were 'evenly stressed'. In the present context, these descriptions are viewed as support for concepts such as isochrony and syllable segregation as a possible core deficit in apraxia of speech.

Statement of the problem

Among other proposed descriptive features of CAS, a reduction in the normal temporal variation of speech events has significant promise as a diagnostic marker. From a conceptual perspective, an impairment in speech timing may be a highly prevalent correlate of apraxia of speech, occurring across differing levels of severity of expression and persisting past the period when other segmental and suprasegmental variables normalize (i.e. consider the advanced ages of the affected

family members included in Piggott and Kessler Robb, 1999). From a statistical perspective, a timing deficit may be highly specific for apraxia of speech compared, for example, to a stress impairment that may be prevalent in other types of communicative disorders. Finally, from a technical perspective, durational data are relatively easy to quantify using acoustic techniques.

The goals of this report are (a) to describe a metric to quantify the variation in speech and pause durations in conversational speech, (b) to provide reference data on this variable in representative samples of 3–6-year-old children with typical and delayed speech, (c) to provide data on the variation in speech and pause events in children who have been described as having isochronous and segregated syllables consistent with apraxia of speech and (d) to assess the potential of the temporal variation metric as a diagnostic marker of childhood apraxia of speech. Following the research and service delivery rationale proposed in Shriberg, Aram and Kwiatkowski (1997a), the term *suspected apraxia of speech* (*sAOS*) is used to denote case studies or participants recruited for research who may or may not be true positives for childhood apraxia of speech.

Method

Participants

Audiocassettes of conversational speech samples from 75 children were obtained from a large audiotape archive. The archive included recorded speech samples and accompanying case records of children who had participated in one or more studies of child speech-sound disorders, either as experimental or as control subjects. Table 1 includes summary descriptive data for children in the following three speaker groups.

Table 1. *Summary descriptive data for three groups of participants*

	N			Age		PCCR[a]		AWU[b]	
Group	Male	Female	Total	*M*	*SD*	*M*	*SD*	*M*	*SD*
Normal Speech Acquisition (NSA)									
3–4 years	5	5	10	3;7	0;4	95.8	2.6	4.5	1.4
4–5 years	5	5	10	4;3	0;2	97.7	1.1	3.9	0.9
5–6 years	5	5	10	5;8	0;2	99.3	0.4	5.3	1.0
Total	15	15	30	4;6	0;11	97.6	2.2	4.6	1.2
Speech Delay (SD)									
3–4 years	5	5	10	3;5	0;3	74.5	11.7	3.0	0.9
4–5 years	5	5	10	4;3	0;3	71.0	10.5	4.3	0.9
5–6 years	5	5	10	5;5	0;3	68.1	11.1	3.9	0.6
Total	15	15	30	4;4	0;11	71.2	11.0	3.7	0.9
Suspected Apraxia of Speech (sAOS)									
3–4 years	2	0	2	3;5	0;4	66.3	3.4	2.6	0.5
4–5 years	1	1	2	4;5	0;7	72.8	3.6	4.3	0.9
5–14 years	10	1	11	8;0	3;3	79.7	11.3	3.4	1.4
Total	13	2	15	6;11	3;4	77.0	10.9	3.4	1.3

[a]Percentage of Consonants Correct-Revised (PCCR; Shriberg, Austin *et al.*, 1997a).
[b]Average Words Per Utterance (AWU; Shriberg *et al.*, 2001).

Normal Speech Acquisition (NSA)

The 30 children classified as having Normal Speech Acquisition (NSA) participated as age-matched controls in prior and ongoing studies of child speech-sound disorders. Children selected for the present study included a total of 15 boys and 15 girls: five boys and five girls each at ages 3, 4 and 5 years. All NSA speakers were classified as having normal speech acquisition using the Speech Disorders Classification System (SDCS) software (Shriberg, Austin, Lewis, McSweeny and Wilson, 1997b; Shriberg, Allen, McSweeny and Wilson, 2001). The SDCS accomplishes classification by comparing a speaker's error profiles to those obtained from age-matched speakers in a reference database. As indicated in the summary data in table 1, all participants with NSA were also required to score above 95% on a severity metric termed the Percentage of Consonants Correct-Revised (PCCR; Shriberg, Austin, Lewis, McSweeny and Wilson, 1997a). This criterion ensured that over 95% of the consonants and vowels/diphthongs attempted in the conversational samples were scored as phonemically correct by one or more transcribers using a common system of narrow phonetic transcription (Shriberg and Kent, 1995). A metric termed Average Words Per Utterance (AWU) included in the PEPPER (Shriberg *et al.*, 2001) output indicated that these 30 children produced an average of 4.6 words per utterance. For children in this age range, AWU correlates in the high 0.90s with mean length of utterance (Shriberg and Kwiatkowski, 1994). These latter data are consistent with the records data indicating that these children had age-appropriate or only mildly delayed language.

Speech Delay (SD)

Conversational samples from 30 3–6-year-old children with speech delay of unknown origin (SD) were selected from the audiocassette archive. As with the samples from the NSA group, these samples were drawn to include also five boys and five girls at ages 3, 4 and 5 years. The inclusionary criteria were that each child met the SDCS criteria for SD. Participants in this group also had to score below 85% on the PCCR metric. As shown in table 1, AWU values for the 30 speakers with SD averaged 3.7.

Suspected Apraxia of Speech (sAOS)

A total of 15 children with speech delay were selected from a subset of 100 children whom speech-language clinicians and/or clinical researchers in collaborative studies had classified as sAOS (Shriberg and McSweeny, 2002). Inclusionary criteria for the present study were based on a set of provisional speech and prosody-voice markers for children with speech delay of unknown origin that were viewed as consistent with sAOS and/or suspected dysarthria (sDYS). An appendix includes the markers and classification procedures used by two experienced research transcribers (the first and fourth authors) to classify the 15 children as sAOS or sDYS.

Candidate participants for the present sAOS speaker group had to meet two criteria. The first inclusionary criterion was that the transcribers perceived the speaker as having a speech-timing deficit consistent with syllable segregation. As listed in the appendix, either of the first two diagnostic markers for sAOS or sDYS satisfied this criterion (notated as Stress/Timing [S/T] in the original procedure; see appendix, table A2). All 15 children selected for the present study were perceived as

having at least three strings of syllable segregation in the section of the conversational speech sample used in the present study (see below). For later use in the data analyses, children with only three such events were subclassified as having *some* syllable segregation; those having four or more strings of segregated sounds were subclassified as having *frequent* syllable segregation.

The second inclusionary criterion for samples meeting the speech-timing criterion was included to increase the level of confidence that each of these participants also had some other behaviours associated with speech motor impairment. Participants had to meet criteria for marginal (1.0–1.5 points) or non-marginal (2 points or more) classification based on their status on the list of speech markers for sAOS and/or sDYS included in the appendix. This list of diagnostic markers had been assembled as part of a prior study (Shriberg and McSweeny, 2002), which was completed before the elaborated list of markers for sAOS proposed in the companion paper in this issue (Shriberg *et al.*, 2003). The outcomes from this procedure were that one participant met criteria for only sAOS, some participants met criteria for only sDYS ($n=8$), some participants met criteria for both classifications (i.e. sAOS *and* sDYS = sAOS_sDYS; $n=2$) and some participants not clearly meeting either classification were termed Indeterminate (i.e. AOS *or* DYS = AOS/DYS; $n=4$).

As indicated in table 1, the participants with sAOS were nearly all male (87%), which is consistent with the reported high proportion of males in studies of sAOS (Hall *et al.*, 1993; Shriberg, Aram and Kwiatkowski, 1997b, 1997c). The average age of the speakers with sAOS was 6 years, 11 months, somewhat older than the average age of speakers in the two comparison groups and also consistent with reports in the literature. Failure to normalize speech delay by 6 years of age is a frequently cited diagnostic characteristic of sAOS, and therefore, studies that ascertain children by clinical referral typically include a substantial number of older children (early elementary-aged to adolescents).

Speech analyses

A subset of 24 utterances from each of the 75 conversational samples was selected using the exclusionary criteria provided in a system of prosody-voice coding (Shriberg, Kwiatkowski and Rasmussen, 1990). These exclusionary criteria, which addressed linguistic, paralinguistic and audio recording requirements, ensured that all utterances were linguistically and technically appropriate for the acoustic analysis.

Digitizing and segmentation

Each of the speakers' 24 utterances was digitized at 44.1 kHz (16 bit quantization). Computer-assisted detection methods were developed for Matlab (version 6.5) to determine the onset and offset of speech and pause events within each of the 24 utterances in each conversational speech sample. For this procedure, the assistant was required to identify a region on a rectified and digitally filtered ($f_{lp}=30$ Hz; zero-phase shift forward and reverse; Butterworth, 8 pole) display of each waveform that contained the largest amplitude 'pause' event. The maximum value of the selected pause region was used to establish a threshold for separating 'speech' from 'pause' events. The onsets and offsets for speech events were defined as values

above the threshold for pause events, as identified in the display for each waveform. Speech and pause durations less than 100 ms, the duration within which stop closures and burst releases tend to occur, were eliminated from the data set. Thus, the pauses remaining were located primarily within and between words, at prosodic phrases and at syntactic boundaries. Figure 1 illustrates these procedures, with the example described in the caption providing additional information on both the procedures and the coefficients of variation described in the next section. The duration values for each eligible speech and pause event were exported to a database.

The coefficient of variation and the coefficient of variation ratio

The auditory percept of isochrony and/or segregated speech, as reviewed previously, includes a lack of normal variation in the temporal aspects of speech. The scope of the present paper does not include extended consideration of factors (other than temporal regularity) that contribute to the percept of isochrony (e.g. Fowler, Smith and Tassinary, 1986; Zellner, 1994). To quantify this percept, a metric was developed to assess a speaker's durational variability within speech events and pause events. The coefficient of variation (CV = SD/Mean) was computed across all of the durations within an utterance for each event type (i.e. speech and pause). This transformation adjusted for changes in event durations that were related to variations in speech rate across utterances and speakers. CV values were averaged across utterances for each subject. Thus, the lower the CV, the lower the variation in the duration of speech events (i.e. isochrony) or pause events in conversational speech.

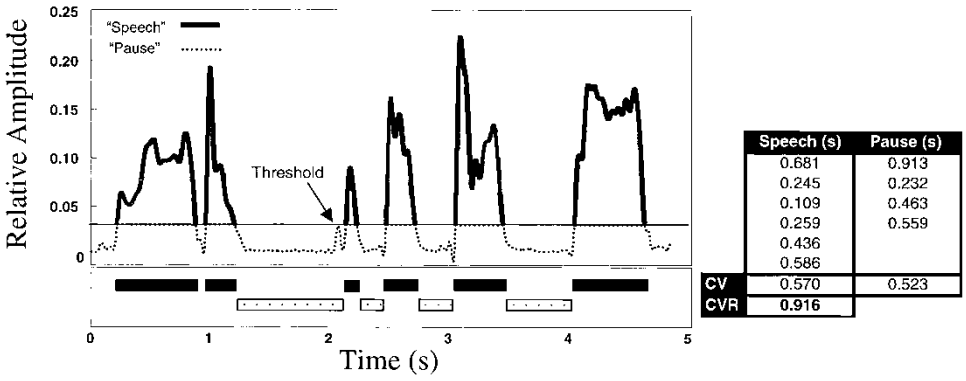

Figure 1. *An example of the measurement procedure used to quantify the duration of speech and pause events for a 5-second utterance. The maximum value of the selected 'pause' region was used to establish an amplitude threshold for separating 'speech' from 'pause' events. Boundaries associated with each pause event were identified as values in the rectified and filtered waveform below the threshold, and speech events were identified as values that were above the threshold. Durations below 100 ms, such as the first pause event in the example waveform, were not included in the analysis. For each utterance, the coefficient of variation (CV) was computed separately for speech and pause events. The coefficient of variation ratio (CVR) for the utterance was computed by dividing the CV of pause events by the CV of speech events. The CVR for the utterance displayed above was 0.916.*

A metric termed the *coefficient of variation ratio* (*CVR*) was calculated to express the relationship between variations in the duration of speech events compared to variation in pause events. Because the CVs for pauses were larger than the CVs for speech, the CVR was calculated by dividing the CV for pause events by the CV for speech events. Thus, low CVRs indicate proportionally less variability in pause events, and high CVRs indicate proportionally less variability in speech events. Figure 1 provides an example of these values as obtained from one 5-second utterance.

Results

Reference data

Table 2 is a summary of the number and distributional characteristics of the durations of speech and pause events recorded for participants in each of the three groups. These values, cross-tabulated by sex and age within each of the speech status groups, provide reference data for typically developing children and children with speech delay. The summary values in table 2 are based on the total number of tokens pooled across speakers. As shown, there were comparable numbers of speech and pause events across the three groups and within the three age levels for each sex. The four descriptive statistics for the distributions of each class of events within each cell suggest that: (a) the means and ranges of speech and pause durations within and across groups were relatively similar; (b) the standard deviations in each cell were fairly large relative to the mean values; (c) the skew values for pause events were larger than those for speech events, the latter of which were more consistent with normal distributions; and (d) the standardized kurtosis values for both events were relatively, but not inordinately, high.

Temporal variation analyses

Description

Table 3 is a summary of the speech CV and pause CV data for participants in each of the three speaker groups. The values for each of the four statistical moments are consistent with normal distributions in each domain for each of the three speaker groups. As with the raw values shown in table 2, these values were not averaged by individual speakers. Rather, they were averaged on the total pool of events indicated in the entries for each row in the column titled Events.

Inferential statistics

Figure 2, panel A is a display of the means and standard deviations for the speech and pause CVs for participants in the three speaker groups. Unlike the data in tables 1 and 2, these CV data were first averaged by speaker and then averaged over the number of speakers in each group. The trends for the speech CVs are consistent with the construct of isochrony in apraxia of speech, with NSA speakers having the highest temporal variation in speech events and sAOS speakers having the least variation. Two-sample t tests using two-tailed significance criteria were completed on these data. Family-wise Bonferoni alpha adjustments were deemed unnecessary, as the six comparisons for each event class could be divided into three logical

Table 2. *Summary of the number and distributional characteristics of the durations (in seconds) of speech and pause events recorded for participants in each of the three groups*

	Demographics			Speech Events					Pause Events				
Group	Age (yrs)	Sex	*n*	*n*	*M*	*SD*	Skew	Kurtosis	*n*	*M*	*SD*	Skew	Kurtosis
Normal Speech Acquisition (NSA)	3	Male	5	319	0.31	0.21	2.05	6.60	204	0.25	0.18	2.26	−0.85
	3	Female	5	354	0.36	0.25	1.85	4.11	296	0.35	0.58	6.14	−0.15
	4	Male	5	403	0.31	0.20	1.84	4.09	258	0.26	0.19	2.76	−0.15
	4	Female	5	502	0.28	0.18	2.42	8.90	268	0.33	0.39	3.54	−0.83
	5	Male	5	439	0.34	0.22	1.43	2.37	332	0.34	0.36	3.16	−0.14
	5	Female	5	528	0.31	0.22	2.57	10.70	332	0.34	0.36	3.16	−0.87
Speech Delay (SD)	3	Male	5	272	0.31	0.22	2.15	5.25	197	0.32	0.31	4.05	0.60
	3	Female	5	309	0.34	0.21	1.47	2.83	188	0.28	0.29	5.03	0.78
	4	Male	5	512	0.31	0.21	1.99	5.05	304	0.25	0.17	2.58	−0.18
	4	Female	5	353	0.31	0.22	2.06	5.19	233	0.30	0.29	3.43	0.14
	5	Male	5	379	0.29	0.18	2.22	8.08	267	0.29	0.28	2.94	−0.07
	5	Female	5	429	0.32	0.21	2.06	7.16	301	0.27	0.23	2.70	0.18
Suspected Apraxia of Speech (sAOS)	All	Both	15	875	0.30	0.21	2.49	9.65	799	0.38	0.39	2.97	−0.72

Table 3. *Summary of the Coefficient of Variation (CV) data for speech and pause events for participants in the three speaker groups*

	Demographics			Speech CV					Pause CV				
Group	Age (yrs)	Sex	*n*	Events	*M*	*SD*	Skew	Kurtosis	Events	*M*	*SD*	Skew	Kurtosis
Normal Speech Acquisition (NSA)	3	M	5	86	0.48	0.24	0.25	0.06	56	0.48	0.30	0.33	−0.87
	3	F	5	78	0.53	0.26	0.32	0.69	66	0.58	0.38	0.76	−0.15
	4	M	5	86	0.46	0.24	0.66	1.07	62	0.43	0.25	0.29	−0.92
	4	F	5	73	0.50	0.23	0.68	0.59	59	0.57	0.35	0.36	−0.83
	5	M	5	86	0.52	0.23	−0.02	−0.32	68	0.64	0.31	0.50	0.14
	5	F	5	86	0.49	0.24	0.29	0.05	64	0.53	0.30	0.18	−0.87
Speech Delay (SD)	3	M	5	84	0.43	0.24	0.20	−0.66	48	0.52	0.31	0.59	0.60
	3	F	5	73	0.46	0.21	−0.11	−0.54	37	0.57	0.37	0.79	0.78
	4	M	5	103	0.50	0.24	0.38	−0.38	69	0.47	0.24	0.41	−0.18
	4	F	5	80	0.45	0.23	0.35	0.08	54	0.43	0.31	0.75	0.14
	5	M	5	77	0.46	0.22	0.15	−0.53	67	0.46	0.33	0.84	−0.07
	5	F	5	89	0.47	0.23	0.33	0.17	70	0.49	0.31	0.72	0.18
Suspected Apraxia of Speech (sAOS)	All	Both	15	207	0.43	0.22	0.24	−0.45	189	0.54	0.32	0.27	−0.71

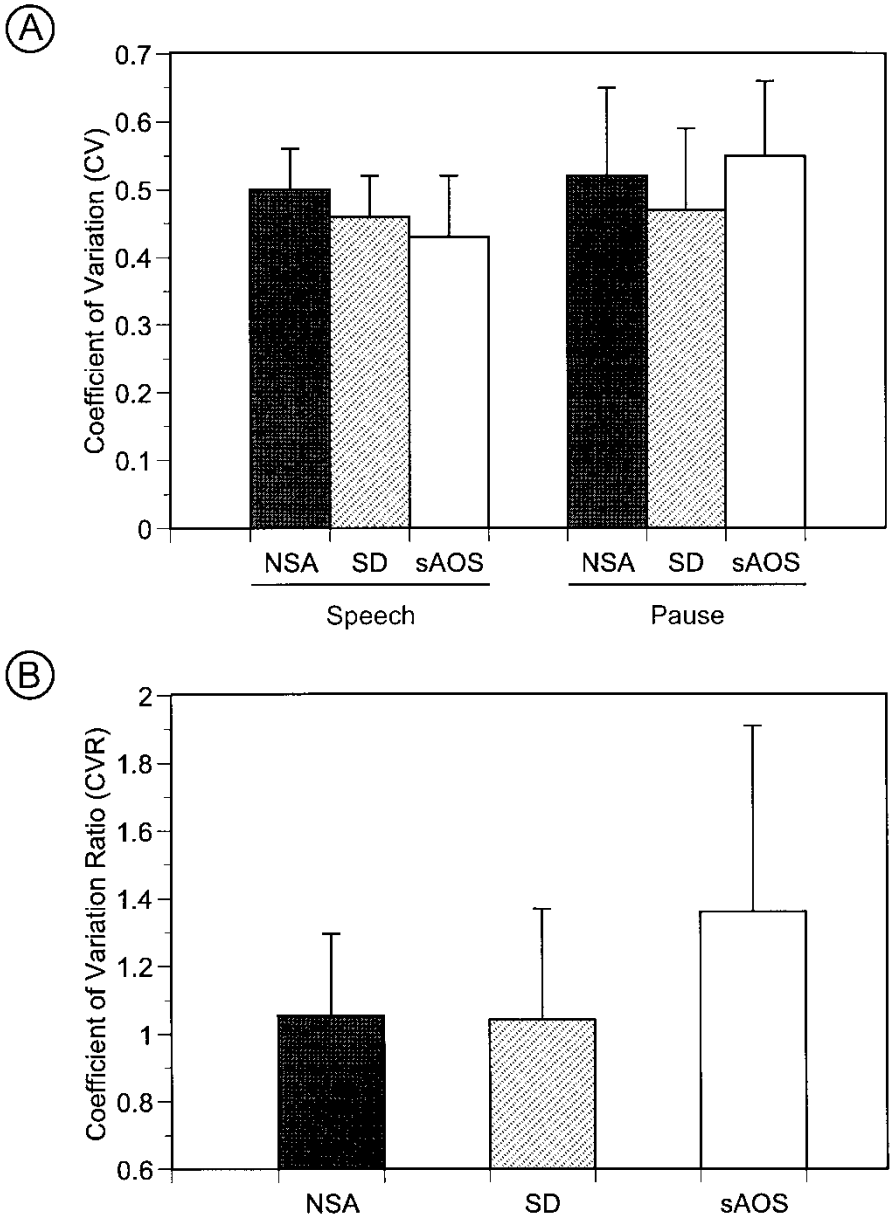

Figure 2. *Panel A is a display of the coefficient of variation (CV) data for speech and pause events for speakers in the three study groups. Panel B is a display of the coefficient of variation ratio data for speech and pause events in each group.*

families. The comparisons of NSA with SD ($t = 2.46$, df $= 58$, $p = 0.017$) and of NSA with sAOS ($t = 3.06$, df $= 43$, $p = 0.0040$) were statistically significant, but the comparison of SD with sAOS ($t = 1.25$, df $= 43$, $p = 0.22$) was not. For the pause CVs in figure 2, panel A, the two reference comparisons of NSA with SD ($t = 1.71$,

df = 58, $p = 0.093$) and of NSA with sAOS ($t = 0.49$, df = 69, $p = 0.39$) were not statistically significant, but the comparison of SD with sAOS speakers ($t = -2.23$, df = 43, $p = 0.031$) on the CV metric was statistically significant.

Figure 2, panel B provides a graphic summary of the CVR data, also averaged first by speaker and then averaged across speakers in each group. Each of the statistical findings for the three comparisons, NSA with SD ($t = 0.17$, df = 53, $p = 0.17$), NSA with sAOS ($t = -2.07$, df = 16, $p = 0.055$) and SD with sAOS ($t = -2.08$, df = 19, $p = 0.052$), exceeded the conventional 0.05 alpha level and thus was not statistically significant.

Effect sizes

Table 4 is a summary of the effect sizes calculated for all between-group comparisons of the CVs for speech and pause events and of the CVR metrics. The effect sizes obtained in the present study were conservatively classified into the four categories shown in the footnote for table 4. As listed in table 4, the effect sizes obtained in this study provide considerable acoustic-based support for the percept of reduced variability of speech and pause events in children with sAOS.

As shown in figure 2 and table 4, the largest effect size was for the speech CVs of speakers with NSA versus sAOS. The sAOS speakers had less variability in their duration of speech events (ES = 0.92). The other large effect size for the CVs was obtained for the comparison of pause CVs for speakers with SD versus sAOS. As shown in figure 2, speakers with sAOS had greater variation in pause durations (ES = 0.71).

There were two large effect sizes obtained for the CVR data. As indicated in table 4 and shown in figure 2, speakers with sAOS had larger CVR ratios (i.e. more variable pause durations and less variable speech durations) than both the speakers with NSA (ES = 0.79) and the speakers with SD (ES = 0.73).

To summarize, the direction and magnitudes of differences in CV and CVR values among groups, as indexed by the generally substantial effect sizes, provide quantitative support for the percept of isochrony in the speakers with sAOS assessed in this study. Findings are also interpreted as support for the sensitivity of

Table 4. *Effect sizes for the between-group comparisons of findings for the Coefficient of Variation (CV) for speech events and pause events and for the Coefficient of Variation Ratio (CVR)*

Variable	Comparison	Effect Size	Magnitude
CV for Speech Events	NSA-SD	0.64	Medium
	NSA-sAOS	0.92	Very Large
	SD-sAOS	0.38	Small
CV for Pause Events	NSA-SD	0.44	Small
	NSA-sAOS	0.23	Small
	SD-sAOS	0.71	Large
CVR	NSA-SD	0.04	–
	NSA-sAOS	0.79	Large
	SD-sAOS	0.73	Large

[a]Effect size adjectives: *Small* (0.20-0.49); *Medium* (0.50-0.69); *Large* (0.70-0.89); *Very Large* (0.90-0.99).
[b]NSA: Normal Speech Acquisition; SD: Speech Delay; sAOS: suspected Apraxia of Speech.

the CVR metric as a potential diagnostic marker of AOS. This latter claim is based on the findings of no appreciable difference in the CVRs of speakers in the NSA versus SD groups (ES = 0.04) but large differences in the NSA-sAOS comparison (ES = 0.79) and in particular, the crucial (i.e. for diagnostic specificity) SD-sAOS comparison (ES = 0.73).

Individual findings

Figure 3 provides information on the distributions of individual CVRs for participants in the three speaker groups. The boxplots provide the median CVR value, a 95% confidence interval around the median value (smaller boxes), the interquartile range (25^{th} and 75^{th} percentiles, larger boxes) and individual CVR data points. The individual data points for the 30 participants in each of the two reference groups are not discernable in this figure. However, it is the spread of scores in each group and the individual CVRs for the 15 children with sAOS that are the focus of figure 3. Four observations about the data in figure 3 warrant comment.

First, the distribution of CVRs for the 15 children with sAOS differs in several ways from the distributions of CVRs for speakers with NSA and SD. There is one speaker with sAOS who had an extremely high CVR (2.76). Further inspection indicated that this speaker had the lowest CV (0.24) for speech events, coupled with the fourth highest CV for pause events. Inspection of the rank-ordered data for the other 4 speakers with high CVRs indicated similar patterns, with CVs for speech events ranked among the lowest and CVs for pause events ranked among the highest within the distributions for the 15 speakers with sAOS.

A second observation is that the elevated median for the speakers with sAOS (in

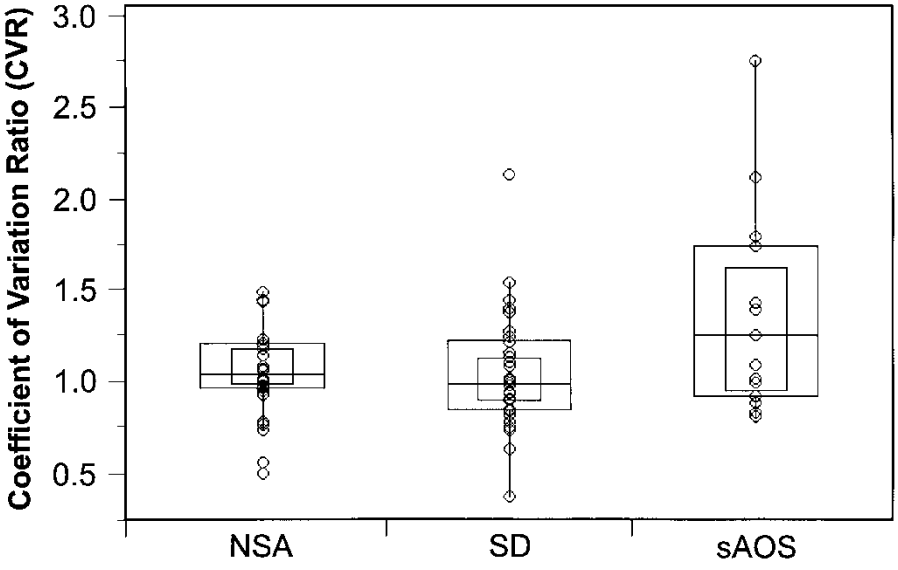

Figure 3. *Boxplots of the coefficient of variation ratio (CVR) data for participants in the three study groups. Each boxplot provides CVR information on the median, confidence interval around the median, the interquartile range and the individual values.*

comparison to the medians for the two reference groups) appears to be associated with the lack of scores in the lower CVR range. Note that the four high CVRs elevate the mean for this group, but extreme scores do not affect the median. Thus, in comparison with the more narrow, balanced ranges of variability of speech and pause events for speakers in the NSA and SD groups—approaching a CVR ratio of 1.0—there is a trend for proportionally more of the speakers with sAOS to have more variation in the durations of their pause events than in their speech events (i.e. their CVRs are above the reference ratios converging on approximately 1.0).

A third observation is that the above findings were not moderated by other characteristics of the sAOS participants. Inspection of the individual data (including use of scatter plots and fitted regression lines) indicated that individual CVs for speech and pause events were unrelated to a participant's etiological classification. Thus, the CVR values of the 7 sAOS participants classified as AOS, AOS_DYS or AOS/DYS (see appendix) were no higher or lower than the CVR values of the 8 participants classified as DYS. Moreover, the magnitudes of the CVs and resulting CVRs were unrelated to the severity of speech involvement (i.e. PCCR values) of the participants with sAOS. Finally, CV and CVR values for the sAOS participants were unrelated to the frequency of perceptually observed syllable segregation in the conversational speech sample (i.e. *some* versus *frequent*).

The fourth observation is viewed as particularly promising for the potential of the CVR metric as a diagnostic marker of AOS. Of the four highest CVRs for sAOS participants shown in figure 3, three are from members of the British family described previously. As reviewed, affected members of the family, including the three participants in the present study, have an orofacial apraxia that appears to be genetically transmitted as an autosomal dominant trait. All affected members also have suspected apraxia of speech. These three participants had the highest, second highest and fourth highest CVR scores (the latter fractionally lower than the third highest, see figure 3) obtained for the 45 participants.

Discussion

The finding that at least some of the speakers with suspected apraxia of speech had notably decreased variation in the duration of their speech events is viewed as support for the perceptual construct of isochrony in these speakers. From a measurement perspective, the CVR findings for the three speakers with an inherited form of apraxia of speech lend a form of construct validity to this proposed diagnostic marker. Some methodological constraints are appropriate to review before additional observations and discussion.

Methodological considerations

There are several procedural considerations that limit generalizations from the findings of this study. An initial consideration is that the decision to include only children whose speech was perceptually consistent with the correlates of isochrony or syllable segregation limits generalizations to participants with these descriptive features of sAOS. The primary goal of the study was to determine if an acoustic correlate of this percept could be developed and documented. It is not clear what the implications of the present findings are for children who may meet other criteria for sAOS but are not perceived as having the characteristics of isochrony and

syllable segregation described in the literature on apraxia of speech and used as inclusionary criteria in the current study.

A second constraint on generalizations from the present findings concerns the validity of the several methodological decisions in the procedures for acoustic analyses. Examples include the selection of the threshold for discriminating speech from non-speech (pause) events and the 100 ms cut-off point for pause events (cf. Hieke, Kowal and O'Connell, 1983; Zellner, 1994). This latter value, for example, could be viewed as too long and wasteful of possibly informative pausal events of shorter duration. Alternatively, inclusion of pausal events above this value may have biased the pause data to appear more variable than was relevant for the present concerns. Cross-validation studies using alternative signal processing approaches and speech-pause criteria are needed for optimal theoretical coherence and diagnostic accuracy.

Finally, the older ages of several of the speakers with sAOS (relative to the data provided in the reference values) may be associated with unknown effects. Although the AWU data provide some control for possible effects on CVR of variables associated with utterance length, there are other age-related variables associated with social and biological correlates of age (e.g. linguistic and craniofacial variables) that may have impact on speech timing. Studies that include larger groups of 3- to approximately 16-year-old children with sAOS are needed to determine whether interpretation of the current findings is constrained by the older ages of some of the speakers with sAOS.

Speech processing perspectives

The quantitative findings for the CVR metric, particularly the substantial effect sizes for the SD-sAOS comparisons, support the potential of the present approach to quantify the perception of isochrony in young speakers. With due appreciation of the need for cross-validation, the following discussion offers some speculative interpretations of the present findings.

Speech variation

The high CVRs obtained by speakers with sAOS were a function of both reduced variation in the duration of speech events and increased variation in the duration of pause events, relative to the CVRs obtained from the two comparison groups. The lowered variation in the duration of speech events would seem to be consistent with the percept of isochrony—which, in turn, is viewed as a deficit consistent with a disorder of speech praxis. From this perspective, reduced or regularized durational variability might be viewed as the obligatory product of a spatiotemporal movement disorder. Note, however, that based primarily on information gained from a tapping test, Alcock *et al.* (2000a; b) speculated that a timing deficit may be the core disorder.

It is difficult to envision how a proposed constraint on the timing of movements might be viewed as support for a locus of apraxia of speech positioned at 'higher' linguistic levels. For example, it is difficult to account for the reduced variation in timing by appeal to such constructs as unstable underlying representations for word forms or sentence formulation deficits affecting the selection and retrieval of words. As reviewed in Shriberg *et al.* (2003), processes of this sort have been proposed as

the possible loci for lexical and sentential stress differences in children with sAOS, including proposals from our own prior studies. From the present perspective, it would seem more plausible that a constraint on temporal variation located at encoding phases of speech planning underlies the lexical and sentential stress deficits observed in children with sAOS.

Pause variation

Speculation on the possible source of the increased variation in the duration of pause events in the children with sAOS is especially challenging. Arguments can be marshalled that this variability reflects constraints at either speech-motor or cognitive-linguistic processing levels. Clearly, some of the pause events could reflect constraints on such task demands as sentence formulation, word retrieval and representational aspects of lexical, sentential or emphatic stress. Because it was beyond the scope of the present study to complete the type of transcript analysis that could provide the needed information on such correspondences in children with sAOS compared to children in the other two groups, the present data are viewed as non-informative on these issues. However, using the criteria of explanatory parsimony, it would be appropriate to infer a common source for both the speech and the pause findings. Such a view would propose that the variable pause lengths reflect planning and programming processes at the level of speech motor control, much as posited for adults with acquired AOS whose premorbid representational forms and access processes are posited to remain intact. Again, the goals and methods used in the current study were not designed to address these issues.

Conclusions

The primary goal of this study was to develop and assess the diagnostic accuracy of a metric termed the Coefficient of Variation Ratio, an acoustically based measure that quantifies the relative durational variability of speech versus pause events. Descriptive statistics for the metric suggest that it meets criteria for interval-level analyses. For the age ranges typically studied in childhood apraxia of speech, additional research is needed to assemble a reference database extending through adolescence. Concurrent validity for the metric was supported by its sensitivity to the speech of three children whose patterns of speech and prosodic errors have been linked to a genotype associated with orofacial apraxia. Findings suggest that these children (and others with sAOS) have reduced temporal variation in speech events but increased temporal variation in pause events. Findings for the reduced variation in speech events were interpreted as consistent with a praxis deficit in speech motor control. However, findings indicating that children with sAOS have greater variability in the durations of pause events, together with findings indicating that such children have deficits in the perception of correct stress, present continued challenges to the explanatory construct of a core praxis deficit in childhood apraxia of speech.

Acknowledgements

Our sincere thanks to the speech-language pathologists at the Children's Hospital of Pittsburgh who helped us identify some of the participants for this study and to the children and their parents who volunteered their time to participate in this project. Thanks to our laboratory colleagues and collaborators in Pittsburgh, Madison and Cleveland for their assistance and contributions at different stages of this project, including Roger Brown, Lisa Freebairn, Sharon Gretz, Sheryl Hall, Heather Karlsson, Barbara Lewis, Connie Nadler, Tammy Nash, Dayna Pitcairn, Carmen Rasmussen, Heather Rusiewicz, Jennine Sprangers, David Wilson and Marie Wirka. Special thanks to Katherina Hauner for significant editorial suggestions and assistance. This research is supported by the National Institute on Deafness and Other Communicative Disorders, NIDCD DC00496.

References

ACKERMANN, H. and HERTRICH, I., 1994, Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica*, **46**, 70–78.

ALCOCK, K. J., PASSINGHAM, R. E., WATKINS, K. E. and VARGHA-KHADEM, F., 2000a, Oral dyspraxia in inherited speech and language impairment and acquired dysphasia. *Brain and Language*, **75**, 17–33.

ALCOCK, K. J., PASSINGHAM, R. E., WATKINS, K. E. and VARGHA-KHADEM, F., 2000b, Pitch and timing abilities in inherited speech and language impairment. *Brain and Language*, **75**, 34–46.

BELTON, E., SALMOND, C. H., WATKINS, K. E., VARGHA-KHADEM, F. and GADIAN, D. G., 2003, Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human Brain Mapping*, **18**, 194–200.

FEE, E. J., 1995, The phonological system of a specifically language-impaired population. *Clinical Linguistics and Phonetics*, **9**, 189–209.

FISHER, S. E., VARGHA-KHADEM, F., WATKINS, K. E., MONACO, A. P. and PEMBREY, M. E., 1998, Localisation of a gene implicated in a severe speech and language disorder. *Nature Genetics*, **18**, 168–170.

FOWLER, C. A., SMITH, M. and TASSINARY, L. G., 1986, Perception of syllable timing by prebabbling infants. *Journal of Acoustical Society of America*, **79**, 814–825.

GOAD, H., 1998, Plurals in SLI: prosodic deficit or morphological deficit? *Language Acquisition*, **7.2–4**, 247–284.

HALL, P. K., 1992, At the center of the controversy: developmental apraxia. *American Journal of Speech-Language Pathology*, **1**, 23–25.

HALL, P. K., JORDAN, L. S. and ROBIN, D. A., 1993, *Developmental Apraxia of Speech: Theory and clinical practice* (Austin, TX: Pro-ed).

HIEKE, A. E., KOWAL, S. and O'CONNELL, D. C., 1983, The trouble with 'articulatory' pauses. *Language and Speech*, **26**, 203–213.

HURST, J. A., BARAITSER, M., AUGER, E., GRAHAM, F. and NORELL, S., 1990, An extended family with a dominantly inherited speech disorder. *Developmental Medicine and Child Neurology*, **32**, 347–355.

KENT, R. D., 2000, Research on speech motor control and its disorders: a review and prospective. *Journal of Communication Disorders*, **33**, 391–428.

KENT, R. D. and ROSENBEK, J. C., 1983, Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research*, **26**, 231–249.

KENT, R. D., WEISMER, G., KENT, J. F., VORPERIAN, H. K. and DUFFY, J. R., 1999, Acoustic studies of dysarthric speech: methods, progress, and potential. *Journal of Communication Disorders*, **32**, 141–180.

LAI, C. S. L., FISHER, S. E., HURST, J. A., LEVY, E. R., HODGSON, S., FOX, M., JEREMIAH, S., POVEY, S., JAMISON, D. C., GREEN, E. D., VARGHA-KHADEM, F. and MONACO, A. P., 2000, The SPCH1 region on human 7q31: genomic characterization of the

critical interval and localization of translocations associated with speech and language disorder. *American Journal of Human Genetics*, **67**, 357–368.

LAI, C. S. L., FISHER, S. E., HURST, J. A., VARGHA-KHADEM, F. and MONACO, A. P., 2001, A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, **413**, 519–523.

MATLAB, 2002, [Computer software] (Natick, MA: The MathWorks, Inc.).

ODELL, K. H. and SHRIBERG, L. D., 2001, Prosody-voice characteristics of children and adults with apraxia of speech. *Clinical Linguistics and Phonetics*, **15**, 275–307.

PIGGOTT, G. L. and KESSLER ROBB, M., 1999, Prosodic features of familial language impairment: constraints on stress assignment. *Folia Phoniatrica et Logopaedica*, **51**, 55–69.

ROBIN, D. A., 1992, Developmental apraxia of speech: just another motor problem. *American Journal of Speech-Language Pathology*, **1**, 19–22.

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

SHRIBERG, L. D., ARAM, D. M. and 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. and KWIATKOWSKI, J., 1997b, Developmental apraxia of speech: II. Toward a diagnostic marker. *Journal of Speech, Language, and Hearing Research*, **40**, 286–312.

SHRIBERG, L. D., ARAM, D. M. and KWIATKOWSKI, J., 1997c, 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. and WILSON, D. L., 1997a, The Percentage of Consonants Correct (PCC) metric: extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, **40**, 708–722.

SHRIBERG, L. D., AUSTIN, D., LEWIS, B. A., MCSWEENY, J. L. and WILSON, D. L., 1997b, The Speech Disorders Classification System (SDCS): extensions and lifespan reference data. *Journal of Speech, Language, and Hearing Research*, **40**, 723–740.

SHRIBERG, L. D., CAMPBELL, T. F., KARLSSON, H. B., BROWN, R. L., MCSWEENY, J. L. and NADLER, C. J., 2003, A diagnostic marker for childhood apraxia of speech: the lexical stress ratio. *Clinical Linguistics and Phonetics*, **17**, 549–574.

SHRIBERG, L. D. and KENT, R. D., 1995, *Clinical Phonetics*, second edition (Boston, MA: Allyn & Bacon).

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

SHRIBERG, L. D., KWIATKOWSKI, J. and RASMUSSEN, C., 1990, *The Prosody-Voice Screening Profile* (Tucson, AZ: Communication Skill Builders).

SHRIBERG, L. D. and MCSWEENY, J. L., 2002, *Classification and misclassification of childhood apraxia of speech*, Tech. Rep. No. 11 (Phonology Project, Waisman Center, University of Wisconsin-Madison).

SHUSTER, L. I., RUSCELLO, D. M. and HAINES, K. B., 1989, *Acoustic patterns of developmental apraxia of speech.* Paper presented at the American Speech-Language-Hearing Association Convention, St. Louis, MO.

VARGHA-KHADEM, F., WATKINS, K. E., ALCOCK, K. J., FLETCHER, P. and PASSINGHAM, R., 1995, Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 930–933.

VARGHA-KHADEM, F., WATKINS, K. E., PRICE, C. J., ASHBURNER, J., ALCOCK, K. J., CONNELLY, A., FRACKOWIAK, R. S., FRISTON, K. J., PEMBREY, M. E., MISHKIN, M., GADIAN, D. G. and PASSINGHAM, R. E., 1998, Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 12695–12700.

WATKINS, K. E., DRONKERS, N. and VARGHA-KHADEM, F., 2002, Behavioural analysis of an inherited speech and language disorder: comparison with an acquired aphasia. *Brain*, **125**, 452–464.

WATKINS, K. E., GADIAN, D. G. and VARGHA-KHADEM, F., 1999, Functional and structural brain abnormalities associated with a genetic disorder of speech and language. *American Journal of Human Genetics*, **65**, 1215–1221.

WHITESIDE, S. P. and VARLEY, R. A., 1998, A reconceptualisation of apraxia of speech: a synthesis of evidence. *Cortex*, **34**, 221–231.

ZELLNER, B., 1994, Pauses and the temporal structure of speech. In E. Keller (ed.), *Fundamentals of Speech Synthesis and Speech Recognition* (Chichester: John Wiley), pp. 41–62.

Appendix: classification procedures

Detailed descriptions of the procedures developed to classify children with suspected apraxia of speech are provided in a technical report (Shriberg and McSweeny, 2002). For the needs of that study, a system was developed to reclassify 100 children referred by speech-language pathologists and research collaborators as having suspected apraxia of speech. The abbreviated description that follows provides only the information needed to understand the classifications as they are discussed in the present study.

Table A1 includes a list of diagnostic markers and their point values for two proposed etiological subtypes of speech delay of currently unknown origin. The Shriberg and McSweeny paper includes markers and classification procedures for a third proposed etiological subtype-speech delay associated with early recurrent otitis media with effusion. For clarity, all information on this third classification has been removed from tables A1 and A2. Only one child in the present study qualified for a multiple category (see below) that involved this latter proposed subtype.

Each marker in the two lists was developed in associated research on sAOS.

Table A1. *Speech and prosody markers and bracketing criteria for two proposed subtypes of speech delay. Participants not meeting criteria for either category are classified as having speech delay of unknown origin (SD)*

		Point Values	
Classification	Behaviour	Bracketed	Non-Bracketed
SD-AOS	1. Excessive/equal stress (Shriberg, Aram *et al.*, 1997b)	0.0	0.0
	2. Inappropriate timing: (syllable segregation; isochrony)	0.0	0.0
	3. Inconsistent errors on the same word type	0.5	1.0
	4. Prearticulatory oral gestures (groping)	1.0	2.0
	5. Postarticulatory repetitions and revisions	1.0	2.0
	6. Metathetic/sequencing errors	1.0	2.0
SD-DYS	1. Excessive/equal stress (Shriberg, Aram *et al.*, 1997b)	0.0	0.0
	2. Inappropriate timing (syllable segregation; isochrony)	0.0	0.0
	3. Slow rate	0.5	1.0
	4. Voice tremor/tremulous	0.5	1.0
	5. Nasal resonance	0.5	1.0
	6. Nasal emission	1.0	2.0
	7. Imprecise vowels/consonants	1.0	2.0

Table A2. *Definitions and criteria used to assign speakers to one of the etiological classification categories. For efficiency, the prefix s (i.e. suspected) is deleted from each classification term*

Classification Category	Abbreviation	Description	Criteria and Symbolization
Speech Delay-Apraxia of Speech	AOS	Speech Delay in association with apraxia of speech	Speakers meet phonological process or residual distortion error criteria for current or prior speech delay, plus have a total of 1–1.5 points (marginal) or 2 points or greater (non-marginal) for the eight segmental and prosody markers for AOS described in table A1. Denote speakers who also meet criteria for either of the two stress/timing markers by the suffix S/T (i.e. AOS_S/T).
Speech Delay-Dysarthria	DYS	Speech Delay in association with dysarthria	Speakers meet phonological process or residual distortion error criteria for current or prior speech delay, plus have a total of 1–1.5 points (marginal) or 2 points or greater (non-marginal) for the six segmental and prosody markers for DYS described in table A1. Denote speakers who also meet criteria for either of the two stress/timing markers for DYS by the suffix S/T (i.e. DYS_S/T).
Speech Delay-Apraxia and Dysarthria	AOS_DYS	Speech Delay in association with apraxia of speech and dysarthria	Speakers meet criteria for both AOS and DYS. Place brackets around stems meeting only marginal criteria for that classification (i.e. AOS_[DYS]). Denote speakers who also meet criteria for either of the two stress/timing markers for AOS or DYS by the suffix S/T (i.e. AOS_DYS_S/T).
Speech Delay-Apraxia of Speech or Dysarthria	AOS/DYS	Speech Delay in association with apraxia of speech and/or dysarthria	Speakers meet only the marginal or non-marginal stress and/or timing criteria for AOS or DYS in table A1 (i.e. they do not meet the segmental criteria for either or both of these classifications).

Bracketed observations, as indicated in the right-most column in table A1, essentially reflect limitations in the frequencies of marker behaviours (i.e. validity and reliability issues). Some speakers met the two-points criterion for both proposed etiological categories, which is termed AOS_DYS (i.e. AOS *and* DYS). Some children met the stress or timing criterion for AOS or DYS, but not the speech criteria that might differentiate the two classifications. Participants meeting such criteria for only the stress or timing marker were classified as AOS/DYS (i.e. AOS *or* DYS). Importantly, all classifications were based on the available audio samples, which were restricted to conversational speech samples for the majority of most participants. As discussed in the Shriberg and McSweeny paper, responses to citation tasks and other types of speech tasks are needed to evoke the types of behaviours associated with apraxia of speech. Thus, the possibility of false positives and false negatives among these classifications is the primary motivation for the term *suspected* apraxia of speech in both studies. It should also be noted that the eight descriptive features for childhood apraxia of speech proposed in Shriberg, Campbell *et al*. (2003) were developed from research completed after the Shriberg and McSweeny (2002) procedures used in the present study.