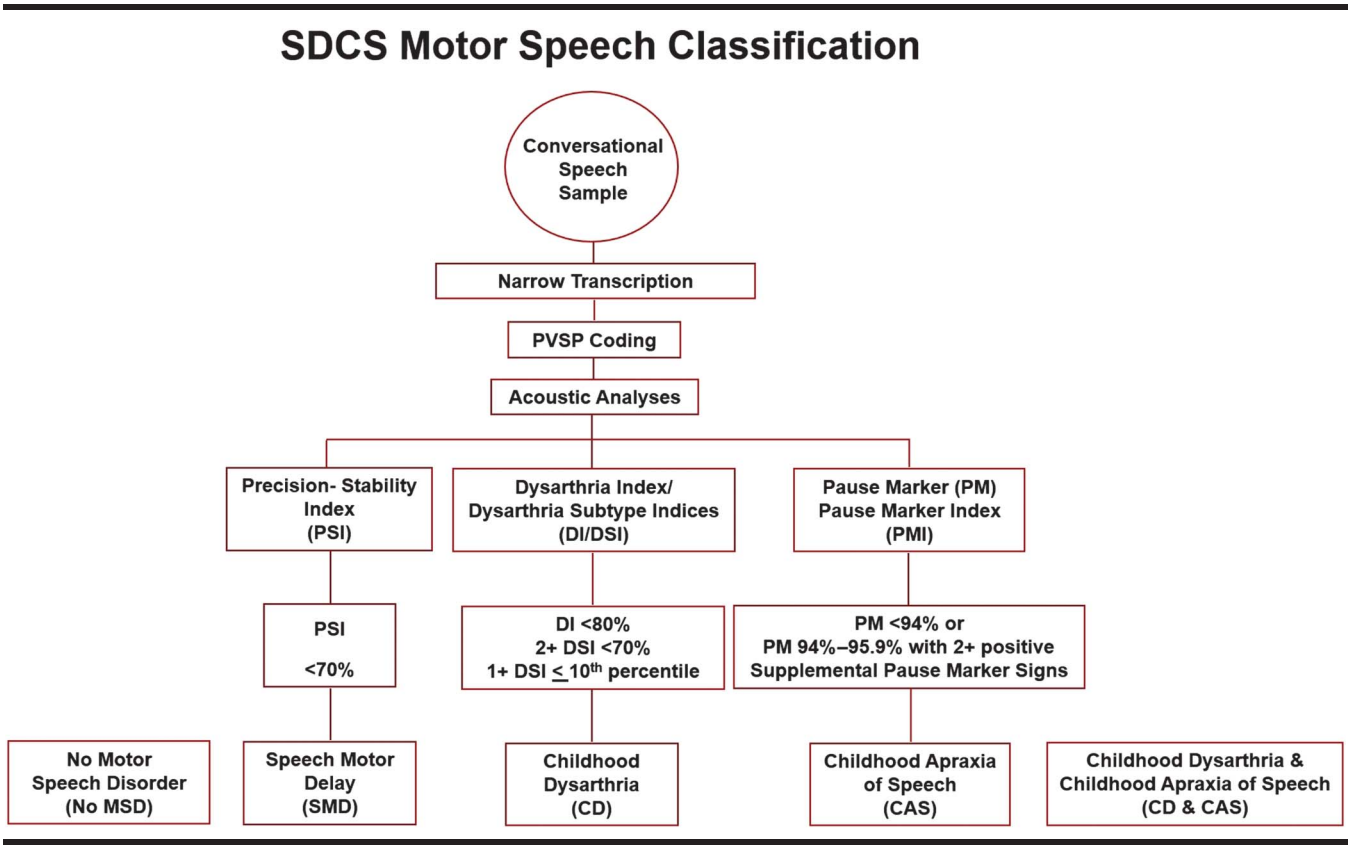
MOTOR SPEECH DISORDERS IN IDIOPATHIC SPEECH DELAY AND IN COMPLEX NEURODEVELOPMENTAL DISORDERS USING THE SPEECH DISORDERS CLASSIFICATION SYSTEM

SUPPLEMENT

This Supplement is a slightly edited version of Appendix B in Baylis and Shriberg (2018). It provides information on the procedure and measures the Speech Disorders Classification System (SDCS) uses to classify a speaker's motor speech status. Figure S1 is a summary of the procedure. Additional information on methods is described in the following sections, in the research articles and technical reports cited in the text, and in Programs to Examine Phonetic and Phonologic Evaluation Records (PEPPER, 2019).

Figure S1. Summary of the procedure used by the Speech Disorders Classification System (SDCS) to classify a speaker's motor speech status. PVSP = Prosody-Voice Screening Profile.



As shown in Figure S1, the procedure begins with an audio-recorded conversational speech sample, obtained following the guidelines described in Programs to Examine Phonetic and Phonologic Evaluation Records (PEPPER; 2019). Next, narrow phonetic transcription, prosody-voice coding, and acoustic analyses are completed following PEPPER guidelines and entered into the SDCS program. Third, the program computes scores on the three measures, as described in the following sections. Last, using the rules for each of the three measures shown in Figure S1 and in the following sections, the program classifies the speaker's motor speech status into one of five classifications. The following sections include descriptions of each measure, including copies of the signs and information on the calculations for each sign in two of the measures.

The Precision–Stability Index

Tables S1 and S2 include the 13 perceptual and 19 acoustic signs of precision and stability of speech that comprise the 32-item Precision–Stability Index (PSI; Shriberg et al., 2010). Table S1 is a facsimile of the PSI scoring sheet, and Table S2 includes information on the calculations completed by the PEPPER (2019) software for each sign. The PSI was developed as a measure to identify and quantify the SDCS classification of motor speech disorders termed speech motor delay. The 32 signs are divided into three domains of speech and seven domains of prosody and voice. Tables S1 and S2 are the PSI forms for individual speakers; outputs are also available that summarize findings from a group of speakers. A technical report includes additional psychometric information and empirical findings for this measure (Mabie & Shriberg, 2017).

As shown in Figure S1, the data for each of the signs are obtained from a standardized procedure to obtain a sample of continuous or conversational speech, including procedures useful for children with low verbal output (McSweeny, 1998; Shriberg & Kwiatkowski, 1985). Phonetic transcription, prosody-voice coding, and acoustic analyses procedures are described and referenced in Shriberg et al. (2010), with additional information in PEPPER (2019). A database of 150 typical speakers used to standardize speech, prosody, and voice signs of speakers 3–17 years of age includes 5 males and 5 females at each age (Potter et al., 2012). A reference database of 50 speakers used to standardize signs for 20- to 80-year-old speakers includes four speakers of the same sex for speakers across each decade from 20 to 69 years of age and five speakers of the same sex from 70 to 79 years of age (Scheer-Cohen et al., 2013).

A z score cutoff of 1.25 is used to divide performance on each sign as within or not within the typical range. Directionality of the z score changes for some signs; for most signs, a minus z score indicates lower precision or lower stability. z scores equal to or less than 1.24 are considered “not positive” for the sign and coded “0.” z scores greater than 1.25 are considered “positive” for the sign and coded “1.” The PSI percentage score is calculated by dividing the number of positive signs by 32 (or less if missing data) and subtracting the quotient from 100, so that lower percentage scores reflect reduced precision and stability. As shown in Figure S1, to be classified as PSI+ (i.e., SMD), a speaker’s PSI score is required to be less than 70%.

The Dysarthria Index and Dysarthria Subtype Indices

Tables S3 and S4 include the 19 perceptual and 15 acoustic signs that comprise the Dysarthria Index and Dysarthria Subtype Indices (DI/DSI). Information for each of the signs is obtained from the same sample of continuous or conversational speech as used for the PSI and processed using the same transcription, prosody-voice, and acoustic methods and reference databases to standardize scores. Table S3 is a facsimile of the DI/DSI form for individual speakers; outputs are also available that summarize findings from a group of speakers. Table S4 includes information on the calculations completed by the PEPPER (2019) software for each sign. The DI/DSI were developed as measures to identify and quantify the SDCS classifications of motor speech disorders termed childhood dysarthria (CD) and CD and childhood apraxia of speech.

The 34 items that operationalize and quantify signs of dysarthria in the DI/DSI were based on the adult-onset signs of dysarthria, including the weightings of each sign (“1” or “2,” bolded) for subtypes of dysarthria, as described in Duffy (2013). A z score cutoff of 1.50 SD units is used to divide performance on each sign as within or not within the typical range. The same procedures used for the PSI scores are used to code each of the 34 DI/DSI items. Also, as described in Mabie and Shriberg (2017), the item weightings for each of the dysarthria subscales shown in Table S3 (which are the same weightings as those in Duffy [2013]) are used to derive percentage scores for each of the five dysarthria subscales. As shown in Figure S1, three criteria must be met to code a speaker as positive for dysarthria. First, the DI score must be less than 80.0%. Second, two or more of the five dysarthria subtypes must have percentage scores less than 70.0%. Last, at least one dysarthria subtype percentile is required to be less than or equal to the 10th percentile (see Mabie & Shriberg, 2017, for how percentile reference data were obtained). Because the same signs are used in as many as three of the five indices, some of the five DSI have high correlation coefficients with one another (Shriberg & Mabie, 2017). Until additional research is available, the descriptive percen-tile data are used only to support the possibility of pure or mixed subtypes of CD.

The Pause Marker and Pause Marker Index

The Pause Marker (PM) is a diagnostic marker to identify speakers with CAS and discriminate CAS from SMD and CD (Shriberg et al., 2017a, 2017b, 2017c, 2017d). As shown in Figure S1, speech data for the PM are obtained from the same sample of conversational speech as used for the PSI and DI/DSI and processed using the same transcription, prosody-voice, and acoustic methods and reference databases to standardize PM scores. There is no PM form to include in this Supplement. As shown in Figure S1, CAS classification requires provisions for several quantitative outcomes.

The PM score is based on the occurrence of four types of inappropriate between-words pauses collectively termed Type I pauses. Tilkens et al. (2017) include reference information for the four Type I pauses as well as for the four inappropriate between-words pauses (Type II pauses) that provide additional information on speech processes for clinical and research applications in CAS.

The PM score is the percentage (subtracted from 100%) of Type I pauses that occur in 24 utterances of a continuous speech sample that meet eligibility criteria for the Prosody-Voice Screening Profile (Shriberg, Kwiatkowski, & Rasmussen, 1990). A minimum of 40 between-words pause opportunities must occur in the speech sample in order to obtain a valid PM score. PM scores lacking 40 between-words pauses are classified as indeterminate. A PM score above 96% (PM−) is classified as within the typical range and the speaker is classified as negative for CAS (CAS−). A PM score below 94% (PM+) is classified as not within the typical range and the speaker is classified as positive for CAS (CAS+).

A PM score from 94% to 95.9% is termed a marginal PM score. To resolve the CAS classification of a speaker with a marginal PM score, findings from three signs termed the supplemental PM signs are used. CAS+ classification of a marginal PM score requires that at least two of the three supplemental PM signs (slow rate, inappropriate stress, transcoding errors) are positive. Marginal PM scores that cannot be resolved by SPMS findings (due to missing data or insufficient pause opportunities) are also classified as indeterminate.

To scale the severity of CAS for clinical and research needs, the Pause Marker Index divides PM+ scores into four ordinal levels (Shriberg et al., 2017d). "Mild" CAS severity scores include PM scores from 90.0% to 93.9%, "mild–moderate" severity scores include PM scores between 85.0% and 89.9%, "moderate–severe" severity scores include PM scores between 80.0%and 84.9%, and "severe" severity scores include PM percentages below 80.0%.

Table S1. Precision–Stability Index (PSI).

PSI: Individual							
	Sign				Sign values		
Linguistic domain	No.	Description	Assessment mode[a]		Value	*z* Score[b]	Code[c]
Vowels			P	A			
	1	Reduced dispersion of corner vowels from center		X			
	2	Reduced dispersion of corner vowels from ^		X			
	3	Reduced average pairwise distance of corner vowels		X			
	4	Increased duration of corner vowels		X			
	5	Increased duration for middle vowels and diphthongs		X			
	6	Reduced % vowel phoneme target consistency	X				
	7	Reduced % vowel target consistency	X				
Consonants							
	8	Reduced % correct glides	X				
	9	Increased relative distortion index: sibilants	X				
	10	Reduced % dentalized sibilants of distorted sibilants	X				
	11	Increased relative distortion index for early consonants	X				
	12	Decreased first moment on /s/ initial singletons		X			
	13	Increased sqrt of the second moment for /s/ initial singletons		X			
	14	Increased sqrt of the second moment for /s/ initial, and /s/ and /z/ final singletons		X			
	15	Increased all consonant–consonant duration		X			
Vowels and consonants							
	16	Increased Diacritic Modification Index (DMI) class: place %	X				
	17	Increased DMI class: duration %	X				
	18	Increased % of epenthesis errors	X				
Phrasing							
	19	Increased PM errors: % of addition, breath, repeat, or long	X				
Rate							
	20	Reduced syllables per second (without pauses)		X			
	21	Increased syllable length in ms (without pauses)		X			
Stress							
	22	Increased % of prosody-voice (PV) 15/16 EE (excessive/equal stress) codes of all coded utterances without fast/acceleration (uncircled and circled)	X				
	23	Increased % of PV15/16 EE codes of all PV15/16 codes (uncircled and circled)	X				

PSI: Individual						
	Sign			Sign values		
Linguistic domain	No.	Description	Assessment mode[a]	Value	*z* Score[b]	Code[c]
Loudness						
	24	Decreased intensity difference, dB fricative + vowel	X			
Pitch						
	25	Decreased F0 for all vowels and diphthongs	X			
	26	Decreased range of characteristic F0 for delimited vowels/diphthongs	X			
Laryngeal quality						
	27	Increased % jitter for vowels[b]	X			
	28	Increased % shimmer for vowels[b]	X			
	29	Decreased HNR dB for vowels	X			
Resonance quality						
	30	Increased % inappropriate resonance	X			
	31	Decreased F1 /a/ (nasal)	X			
	32	Decreased F2 for high vowels (nasopharyngeal)	X			
No. of positive signs:						
No. of signs with data:						
Average sign *z* score:						
Signs score:						

Note. PM = Pause Marker. Sqrt = square root; HNR = harmonics-to-noise ratio.

[a]A = acoustic; P = perceptual. [b]*z* scores referenced to age–sex matched typically developing speakers (Potter et al., 2012; Scheer-Cohen et al., 2013). [c]Code: 0 = *not positive on variable*; 1 = *positive on variable* (*z* score ≤ -1.25). Z scores reversed for increased.

Table S2. Precision–Stability Index Sign definitions.

No.	Sign	Mode[a]		Calculation
		P	A	
1	Reduced dispersion of corner vowels from center		X	There are four corner vowels. The center is defined using the average first and second formant frequencies over the four corner vowels. Dispersion is the weighted mean of the corner vowels of the distance to that center. "Weighted" means each vowel occurrence is separately included in the dispersion calculation. In comparison, the center location calculation is unweighted. The average formant frequency pairs are separately calculated for each of the four vowels. The resulting four frequency pairs are then averaged to get the vowel center.
2	Reduced dispersion of corner vowels from ^		X	The location of any vowel is the average first and second formant frequencies. Dispersion is the average distance of the location of each of the four corner vowels to the location of four.
3	Reduced average pairwise distance of corner vowels		X	This is the average distance from the location of each corner vowel to the location of the other three corner vowels.
4	Increased duration of corner vowels		X	The weighted mean of the length of the corner vowels. "Weighted" means each vowel occurrence is separately included in the calculation.
5	Increased duration of middle vowels and diphthongs		X	This includes six monophthongs and five diphthongs.

No.	Sign	Mode[a] P	Mode[a] A	Calculation
6	Reduced % vowel phoneme target consistency	X		A type is a distinct Y-line word considering just the phonemes (see PEPPER [2019] for a description of X, Y, and Z lines). A token is a specific occurrence. A type can have many tokens. Tokens that have anything questionable in the X or Y lines are ignored. Phonemes that are questionable in the Z line are ignored. Cases where a phoneme occurs twice (e.g., "b" in "bob") count as two types. For a given phoneme in a type, consistency looks at just the corresponding position in the tokens for errors. Within a type, consistency is defined as the probability, for the selected phoneme, that any two randomly selected (without replacement) tokens have the same obtained result (considering either phoneme only or phoneme and diacritics). For error consistency, the two tokens are selected just from those with errors; for target consistency, they are selected from all the tokens (assuming at least one token has an error). For example, if three different obtained results occur with frequencies of I, J, and K, then this probability is: $$\frac{I*(I-1)+J*(J-1)+K*(K-1)}{N*(N-1)}$$ where $N = I + J + K$. To combine these probabilities over types, we weight each probability by $N - 1$, because a type with only one eligible token gives us no information. For our "numerator," we store the sum of each type's probability times its $N - 1$. For our "denominator," we store the sum of each type's $N - 1$. Target consistency considers only those types with at least two tokens where at least one has an error. Phoneme consistency considers just substitutions and deletions to be errors.
7	Reduced % vowel target consistency	X		Complete consistency considers substitutions, deletions, and distortions to be errors. Distortions are diacritics on the Z line only (produced but not intended) aside from stress and juncture diacritics.
8	Reduced % correct glides	X		There are two English glides.
9	Increased relative distortion index: sibilants	X		Percentage of sibilant distortions of all sibilant errors (distortions, substitutions, and deletions).
10	Reduced % dentalized sibilants of distorted sibilants	X		Distortions are diacritics on the Z line only (produced but not intended) aside from stress and juncture diacritics. There are three English sibilants.
11	Increased relative distortion index for early consonants	X		Percentage of distorted early eight consonants of all early eight errors.
12	Decreased first moment on /s/ initial singletons		X	See centroid definition at https://en.wikipedia.org/wiki/Spectral_centroid
13	Increased sqrt of the second moment for /s/ initial singletons		X	Sqrt is the abbreviation for square root.
14	Increased sqrt of the second moment for /s/ initial and /s/ and /z/ final singletons		X	The same as the previous item except that final /s, z/ are included.
15	Increased all consonant–consonant duration		X	Average length in milliseconds of all consonant pairs where the consonants are less than 0.1 s apart and they are not the same consonant or a cognate.
16	Increased Diacritic Modification Index (DMI) class: place %	X		Percentage of phonemes with one or more tongue configuration or position diacritics.
17	Increased DMI class: duration %	X		Percentage of phonemes lengthened or shortened.
18	Increased % of epenthesis errors	X		Percent of epenthesis errors (vowel addition) by token (by word).
19	Increased PM errors: % of addition, breath, repeat, or long	X		Percentage of pause opportunities with one or more of addition, breath, repeat, or long. (Counted even if grope, change, or abrupt is also present.)
20	Reduced syllables per second (without pauses)		X	Syllables per second for first 12 coded utterances after pauses are removed.

No.	Sign	Mode[a] P	Mode[a] A	Calculation
21	Increased syllable length in ms (without pauses)		X	Uses the first 12 coded utterances.
22	Increased % of prosody-voice (PV) codes 15/16 EE codes of all coded utterances without fast/acceleration (uncircled and circled)	X		EE is the abbreviation for excessive/equal stress, an inappropriate stress pattern that can occur on utterances that have PVSP inappropriate code of 15 or 16 (see pp. 31–32 of the PVSP manual [Shriberg, Kwiatkowski, & Rasmussen, 1990]). Inappropriate fast and/or accelerated speech (PV11/12) is defined as greater than four syllables per second for children and greater than six syllables per second for adolescents and adults. Uncircled and circled are treated as inappropriate and appropriate, respectively. Circled codes give speakers the benefit of the doubt when a coding decision is difficult to make.
23	Increased % of prosody-voice codes 15/16 EE codes of all PV15/16 codes (uncircled and circled)	X		The same as above except the denominator is the number of PV15/16 codes of any kind.
24	Decreased intensity difference, dB, fricative + vowel		X	For a fricative–vowel pair, the intensity difference is the intensity of the vowel in dB minus the intensity of the fricative in dB. This uses the average intensity difference over all fricative–vowel pairs in the transcript where both phonemes have been delimited during the acoustic analysis.
25	Decreased F0 for all vowels and diphthongs		X	F0 is the fundamental frequency at the characteristic point for those vowels and diphthongs that were delimited during the acoustic analysis.
26	Decreased range of characteristic F0 for delimited vowels/ diphthongs		X	This is the overall maximum F0 minus the overall minimum F0.
27	Increased % jitter for vowels[b]		X	"Jitter is the cycle-to-cycle variation of fundamental frequency, i.e., the average absolute difference between consecutive periods."
28	Increased % shimmer for vowels[b]		X	"Shimmer (dB) is expressed as the variability of the peak-to-peak amplitude in decibels, i.e., the average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20."
29	Decreased HNR dB for vowels		X	TF32 (Milenkovic, 2001[c]) calculates the SNR (signal-to-noise ratio) in dB. To calculate the HNR (harmonics-to-noise ratio): Power = 10 to the (SNR/10) power HNR = 10 * log10(Power – 1)
30	Increased % inappropriate resonance	X		Inappropriate resonance in the PVSP includes inappropriate codes 30, 31, and 32 (nasal, denasal, and nasopharyngeal)
31	Decreased F1 /a/ (nasal)		X	First formant frequency for /a/
32	Decreased F2 for high vowels (nasopharyngeal)		X	Second formant frequency for /i/ and /u/

Note. PVSP = Prosody-Voice Screening Profile; PM = Pause Marker.

[a]A = acoustic; P = perceptual. [b]Jitter and shimmer definitions adapted from "Jitter and Shimmer Measurements for Speaker Recognition," by M. Farrús, J. Hernando, and P. Ejarque, 2007, *Proceedings of the Interspeech*, 778–781. [c]*TF32*: Department of Electrical and Computer Engineering, University of Wisconsin-Madison.

Table S3. Dysarthria Index (DI) and the five Dysarthria Subtype Indices (DSI).

DI and DSI: Individual											
			Assessment mode[a]		DI sign values		Five DSI[b]				
Linguistic domain	Sign no.	Description	P	A	*z* Score[c]	Code[d]	Ataxia	Spastic	Hyperkinetic	Hypokinetic	Flaccid
Vowels											
	1	Increased percentage of vowels/ diphthongs distortions	X				**X (2)**		**X (2)**		
Consonants											
	2	Number of nasal emissions	X								**X (2)**
	3	Increased percentage of weak consonants	X								X (1)
Vowels and consonants											
	4	Increased Diacritic Modification Index class duration	X				X (1)		X (1)		
Phrasing											
	5	Increased slow/pause time	X						X (1)	**X (2)**	
Rate											
	6	Increased slow articulation/pause time	X				X (1)	**X (2)**	X (1)		
	7	Decreased average syllable speaking rate (with pauses)		X			X (1)	**X (2)**	X (1)		
	8	Decreased average syllable articulation rate (without pauses)		X			X (1)	**X (2)**	X (1)		
	9	Increased fast rate	X							**X (2)**	
	10	Decreased stability of syllable speaking rate		X					X (1)	**X (2)**	
Stress											
	11	Increased excessive/equal/misplaced stress	X				**X (2)**	X (1)			
	12	Increased reduced/equal stress	X							**X (2)**	
Loudness											
	13	Decreased stability of Speech Intensity Index		X			**X (2)**		**X (2)**		
	14	Increased stability of Speech Intensity Index		X				X (1)		**X (2)**	X (1)
	15	Increased soft	X							**X (2)**	X (1)
	16	Decreased Speech Intensity Index		X						**X (2)**	X (1)
Pitch											
	17	Increased low pitch/glottal fry	X					**X (2)**	X (1)		
	18	Increased low pitch	X					**X (2)**	X (1)		
	19	Decreased F0 for all vowels and diphthongs		X				**X (2)**	X (1)		
	20	Decreased range of char. F0 among vowels		X				X (1)	X (1)	**X (2)**	X (1)
	21	Decreased stability of F0 for all vowels and diphthongs		X			X (1)				

DI and DSI: Individual

Linguistic domain	Sign no.	Description	Assessment mode[a]		DI sign values		Five DSI[b]				
			P	A	*z* Score[c]	Code[d]	Ataxia	Spastic	Hyperkinetic	Hypokinetic	Flaccid
Laryngeal quality											
	22	Increased breathy	X							X (1)	**X (2)**
	23	Increased rough	X					X (1)	X (1)		
	24	Increased strained	X					X (1)	X (1)		
	25	Number of utterances with [TREM] (tremulous) comment	X						X (1)		
	26	Increased break/shift/tremulous	X					**X (2)**	X (1)		
	27	Increased multiple features	X					**X (2)**	**X (2)**		
	28	Number of diplophonia	X								**X (2)**
	29	Increased % jitter for vowels		X			X (1)				
	30	Decreased stability of jitter for vowels		X			X (1)				
	31	Increased % shimmer for vowels		X			X (1)				
	32	Decreased stability of shimmer for vowels		X			X (1)				
Resonance quality											
	33	Increased nasal	X					X (1)	X (1)	X (1)	**X (2)**
	34	Decreased F1 for /a/ (nasal)		X				X (1)	X (1)	X (1)	**X (2)**
		Unweighted total possible points					12	15	19	11	10
		Weighted total possible points					15	23	22	19	15

DI summary					**DSI summary**				
No. of positive signs					**Ataxia**	**Spastic**	**Hyperkinetic**	**Hypokinetic**	**Flaccid**
No. of signs with data:		Percentage of positive signs:							
Average sign *z* score DI:		DSI (% nonpositive weighted):							
(% nonpositive signs):		DSI percentile score:							

[a]A = Acoustic; P = perceptual. [b]Very frequent: 80.0%–100%; frequent: 60.0%–79.9%; somewhat frequent: 40.0%–59.9%; somewhat infrequent: 20.0%–39.9%; infrequent: 0.0%–19.9%. [c]*z* scores referenced to age–sex matched, typically developing speakers (Potter et al., 2012; Scheer-Cohen et al., 2013). For the three "Number of" items (2, 25, and 28), this column has a count rather than a *z* score. [d]Code: 0 = *not positive on variable*; 1 = *positive on variable* (*z* score $\leq$ 1.50 or "Number of" $\geq$ 2).

Table S4. Dysarthria Index and the five Dysarthria Subtype Indices' sign definitions.

No.	Sign	Mode[a]		Calculation
		P	A	
1	Increased percentage of vowel/diphthong distortions	X		Distortions are diacritics on the Z line only (produced but not intended) aside from stress and juncture diacritics (see PEPPER [2019] for a description of X, Y, and Z lines).
2	Number of nasal emissions	X		Note that this is a count rather than a percentage. Two or more is coded as significant.
3	Increased percentage of weak consonants	X		The "check" diacritic. The "check" diacritic is used to indicate a weakly produced consonant.
4	Increased diacritic modification index class duration	X		The lengthened diacritic (:) or shortened diacritic (>).
5	Increased slow/pause time	X		Rate of less than two syllables per second due to long pause time only (prosody rate code 10).
6	Increased slow articulation/pause time	X		Rate of less than two syllables per second due to slow articulation and pause time (prosody rate code 9).
7	Decreased average syllable speaking rate (with pauses)		X	The average over the first 12 coded utterances given in the Prosody-Voice Screening Profile (PVSP) log of the number of syllables divided by the duration of the utterance as given in the acoustic analysis.
8	Decreased average syllable articulation rate (without pauses)		X	The average over the first 12 coded utterances of the number of syllables divided by the duration of the utterances, less the pause time.
9	Increased fast rate	X		Rate greater than four syllables per second for children or greater than six syllables per second for adolescents/adults (prosody rate code 11).
10	Decreased stability of syllable speaking rate		X	For any measure that occurs multiple times within a source, the stability of that measure can be calculated: Stability = 100*(1-StanDev/Mean) where "*" indicates multiplication, "/" indicates division, and "StanDev" is the standard deviation. If an item is completely stable, StanDev is zero and stability is 100. As StanDev increases as a fraction of the mean, the stability value drops.
11	Increased excessive/equal/misplaced stress	X		Inappropriate stress that is excessive/equal or misplaced (prosody stress code 15; see pp. 31–32 of the PVSP manual [Shriberg, Kwiatkowski, & Rasmussen, 1990]).
12	Increased reduced/equal stress	X		Inappropriate reduction of stress in stressed syllables, plus lack of appropriate stress variation (prosody stress code 14).
13	Decreased stability of Speech Intensity Index (SII)		X	The SII quantifies the difference in dB between a stop or fricative and the following vowel.
14	Increased stability of SII		X	The SII quantifies the difference in dB between a stop or fricative and the following vowel.
15	Increased soft	X		Inappropriate soft voice; judged to be socially unacceptable in face-to-face communication (voice loudness code 17).
16	Decreased SII		X	See Items 13 and 14 for SII definition relative to the stability of SII.
17	Increased low pitch/glottal fry	X		Inappropriate low-pitched, periodic "popping" voice quality distributed across an utterance (voice pitch code 19).
18	Increased low pitch	X		Pitch is inappropriately low for the speaker's age or gender (voice pitch code 20).
19	Decreased F0 for all vowels and diphthongs		X	F0 is the fundamental frequency at the characteristic point for vowels and diphthongs delimited in the acoustic analysis.
20	Decreased range of char. F0 among vowels		X	This is the overall maximum F0 minus the overall minimum F0.
21	Decreased stability of F0 for all vowels and diphthongs		X	See Item 10 for the definition of stability.

No.	Sign	Mode[a] P	Mode[a] A	Calculation
22	Increased breathy	X		Inappropriate laryngeal quality with insufficient vocal tone relative to unvoiced airflow (voice laryngeal code 23).
23	Increased rough	X		Inappropriate laryngeal quality with an aperiodic "gravelly" sound (voice laryngeal code 24).
24	Increased strained	X		Inappropriate laryngeal quality with a strident, tense sounding vocal tone (voice laryngeal code 25).
25	Number of utterances with [TREM] (tremulous) comment	X		Note that this is a count rather than a percentage. Two or more is coded as significant.
26	Increased break/shift/tremulous	X		Occurrence of a voice break, a pitch shift, and/or a tremulous vowel (voice laryngeal code 26).
27	Increased multiple features	X		Inappropriate co-occurring laryngeal features not covered under one laryngeal code (voice laryngeal code 29).
28	Number of diplophonia	X		Note that this is a count rather than a percentage. Two or more is significant (voice laryngeal code 28).
29	Increased % jitter for vowels[b]		X	"Jitter is the cycle-to-cycle variation of fundamental frequency, i.e., the average absolute difference between consecutive periods."
30	Decreased stability of jitter for vowels		X	See Item 10 for the definition of stability.
31	Increased % shimmer for vowels[b]		X	"Shimmer (dB) is expressed as the variability of the peak-to-peak amplitude in decibels, i.e., the average absolute base-10 logarithm of the difference between the amplitudes of consecutive periods, multiplied by 20."
32	Decreased stability of shimmer for vowels		X	See Item 10 for the definition of stability.
33	Increased nasal	X		Inappropriate excess nasality in assimilative and/or assimilative nasality contexts (voice resonance code 30).
34	Decreased F1 for /a/ (nasal)		X	First formant frequency for /a/.

[a]A = acoustic; P = perceptual. [b]Jitter and shimmer definitions adapted from "Jitter and Shimmer Measurements for Speaker Recognition," by M. Farrús, J. Hernando, and P. Ejarque, 2007, *Proceedings of the Interspeech*, 778–781.

REFERENCES

Baylis, A. L., & Shriberg, L. D. (2018). Estimates of the prevalence of speech and motor speech disorders in youth with 22q11.2 Deletion syndrome. *American Journal of Speech-Language Pathology.* Advance online publication. doi: 10.1044/2018_AJSLP-18-0037.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Mosby.

Farrús, M., Hernando, J., & Ejarque, P. (2007). Jitter and shimmer measurements for speaker recognition. *Proceedings of the Interspeech*, 778-781.

Mabie, H. L., & Shriberg, L. D. (2017). *Speech and motor speech measures and reference data for the Speech Disorders Classification System (SDCS).* (Technical Report No. 23). Phonology Project, Madison, WI: Waisman Center, University of Wisconsin–Madison. Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/

McSweeny, J. (1998). *Procedures to obtain extended conversational speech samples for prosody-voice analysis* (Tech. Rep. No. 7). Phonology Project, Waisman Center, University of Wisconsin-Madison. Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/

Milenkovic (2001). *TF32*. Department of Electrical and Computer Engineering, University of Wisconsin, Madison.

Potter, N. L., Hall, S., Karlsson, H. B., Fourakis, M., Lohmeier, H. L., McSweeny, J. L., . . . Shriberg, L. D. (2012). *Reference Data for the Madison Speech Assessment Protocol (MSAP): A Database of 150 Participants 3-to-18 Years of Age with Typical Speech.* (Tech. Rep. No. 18). Phonology Project, Waisman Center, University of Wisconsin-Madison. Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/

Programs to Examine Phonetic and Phonologic Evaluation Records [PEPPER: Computer software]. (2019). Madison, WI: Waisman Center, University of Wisconsin–Madison.

Scheer-Cohen, A. R., Holt, A. S., Karlsson, H. B., Mabie, H. L., McSweeny, J. L., Tilkens, C. M., & Shriberg, L. D. (2013). *Reference Data for the Madison Speech Assessment Protocol (MSAP): A Database of Fifty 20-to-80 year old Participants with Typical Speech.* (Tech. Rep. No. 20). Phonology Project, Waisman Center, University of Wisconsin-Madison. Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/

Shriberg, L. D., Fourakis, M., Hall, S., Karlsson, H. B., Lohmeier, H. L, McSweeny, J. L., . . . Wilson, D. L. (2010). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics & Phonetics*, *24*, 795-824.

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

Shriberg, L. D., & Mabie, H. L. (2017). *Speech and motor speech classification findings for participants with complex neurodevelopmental disorders.* (Tech. Rep. No. 24). Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., . . . Wilson, D. L. (2017a). A diagnostic marker to discriminate of Childhood Apraxia of Speech from Speech Delay: I. Development and description of the Pause Marker. *Journal of Speech, Language, and Hearing Research, 60,* S1096-S1117.

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., . . . Wilson, D. L. (2017b). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: II. Validity studies of the Pause Marker. *Journal of Speech, Language, and Hearing Research, 60,* S1118-S1134.

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., . . . Wilson, D. L. (2017c). A diagnostic marker to discriminate childhood apraxia of speech from speech delay: III. Theoretical coherence of the Pause Marker with speech processing deficits in Childhood Apraxia of Speech. *Journal of Speech, Language, and Hearing Research, 60,* S1135-S1152.

Shriberg, L. D., Strand, E. A., Fourakis, M., Jakielski, K. J., Hall, S. D., Karlsson, H. B., . . . Wilson, D. L. (2017d). A diagnostic marker to discriminate of Childhood Apraxia of Speech from Speech Delay: IV. The Pause Marker Index. *Journal of Speech, Language, and Hearing Research, 60,* S1153-S1169.

Tilkens, C. M., Karlsson, H. B., Fourakis, M., Hall, S. D., Mabie, H. L, McSweeny, J. L., Shriberg, L.D. (2017). *A diagnostic marker of Childhood Apraxia of Speech.* (Tech. Rep. No. 22). Retrieved from Phonology Project website: http://www.waisman.wisc.edu/phonology/