
Quantifying Speech Rhythm Abnormalities in the Dysarthrias

Julie M. Liss

Arizona State University, Tempe

Laurence White

Sven L. Mattys

University of Bristol, Bristol, United Kingdom

Kaitlin Lansford

Arizona State University

Andrew J. Lotto

University of Arizona, Tucson

Stephanie M. Spitzer

Arizona State University

John N. Caviness

Mayo Clinic Arizona, Scottsdale

Purpose: In this study, the authors examined whether rhythm metrics capable of distinguishing languages with high and low temporal stress contrast also can distinguish among control and dysarthric speakers of American English with perceptually distinct rhythm patterns.

Methods: Acoustic measures of vocalic and consonantal segment durations were obtained for speech samples from 55 speakers across 5 groups (hypokinetic, hyperkinetic, flaccid-spastic, ataxic dysarthrias, and controls). Segment durations were used to calculate standard and new rhythm metrics. Discriminant function analyses (DFAs) were used to determine which sets of predictor variables (rhythm metrics) best discriminated between groups (control vs. dysarthrias; and among the 4 dysarthrias). A cross-validation method was used to test the robustness of each original DFA.

Results: The majority of classification functions were more than 80% successful in classifying speakers into their appropriate group. New metrics that combined successive vocalic and consonantal segments emerged as important predictor variables. DFAs pitting each dysarthria group against the combined others resulted in unique constellations of predictor variables that yielded high levels of classification accuracy.

Conclusions: This study confirms the ability of rhythm metrics to distinguish control speech from dysarthrias and to discriminate dysarthria subtypes. Rhythm metrics show promise for use as a rational and objective clinical tool.

KEY WORDS: dysarthria, speech rhythm, rhythm metrics

A consensus view about what constitutes speech rhythm is singularly lacking. Applied to the characterization of different languages, *rhythm* has been used to refer to the perceptually distinctive alternation of stressed and unstressed syllables—what we may qualify as *contrastive rhythm*. Although multiple acoustic factors are likely to contribute to the perception of differences in contrastive rhythm, much work has focused on the particular role of timing. For example, it has been claimed that Spanish sounds as though successive syllables have similar durations, whether stressed or unstressed (“syllable-timed” or “machine-gun rhythm”), whereas English is perceived as having a high durational contrast between stressed and unstressed syllables (“stress-timed” or “Morse code rhythm”; Lloyd James, 1940; Pike, 1945).

The perception of rhythmic differences between and within languages may be seen as arising both from constraints imposed on the production of the syllable stream by the phonological properties of the languages and by the articulatory implementation of these constraints. Following Dauer (1983), most recent research on speech rhythm has focused on the former—in particular, on cross-linguistic differences in syllable structures. For example, Romance languages such as French, Italian, and Spanish are made up predominantly of open (CV) syllables, whereas Germanic languages

such as Dutch, English, and German tend to have more consonant clusters (CCVC, CCVCC, CCCVC, etc.), particularly in stressed syllables. In addition, unstressed vowels in Germanic languages are typically reduced forms, much shorter than stressed vowels, whereas Romance languages show relatively minimal reduction and shortening of unstressed vowels (although this is subject to a great degree of variation within and between languages). Thus, the greater temporal discrepancy between stressed and unstressed syllables in Germanic languages gives rise to the perceptual rhythmic contrast.

Rhythm abnormalities are common in the dysarthrias. However, rather than arising from phonological constraints, it is at the level of articulatory implementation that motor disorders have their impact on the emergent flow of the syllable stream and on the perceived rhythm of speech. Darley, Aronson, and Brown (1969) applied terms such as *excess* and *equal stress*, *reduced stress*, *short rushes of speech*, and *prolonged segments*, among others, to capture the perceptual experience of rhythmic disturbance. Moreover, they suggested that recognizing the pattern of rhythmic disturbance assists in differential diagnosis—that is, the nature of the underlying pathophysiology can give rise to similar sounding rhythm abnormalities within an etiology group. For example, hypokinetic dysarthria of Parkinson's disease (PD) can sound like “rapid fire” with short rushes of speech; hyperkinetic dysarthria of Huntington's chorea can be wildly irregular and unpredictable in its rhythm; mixed flaccid-spastic dysarthria of amyotrophic lateral sclerosis (ALS) can have an equal and even rhythm, with excessive sound prolongations.

Because of this reported perceptual distinctiveness, quantification of rhythmic patterns should be a productive means of distinguishing among the dysarthrias. To our knowledge, however, few such studies have been published. Indeed, even attempts to acoustically characterize the disordered rhythm within single dysarthria types have met with limited success (Kent & Kim, 2003). The most frequently studied disordered rhythm pattern is that of ataxic dysarthria, with its perceptually even, syllable-by-syllable “scanning” character. Like the original studies that failed to find syllabic isochrony in Spanish (Borzone de Manrique & Signorini, 1983) and in French (Wenk & Wiolland, 1982), syllabic isochrony was not found to underlie the perception of scanning speech in either ataxic German (Ackermann & Hertrich, 1994) or ataxic American English (Kent, Kent, Rosenbek, Vorperian, & Weismer, 1997). However, other metrics, particularly those that incorporate vowel duration measures, have shown evidence for a tendency toward more regular vowel durations in ataxic dysarthria than in healthy speakers (Hartelius, Runmarker, & Andersen, 2000; Henrich, Lowit, Schalling, & Mennen, 2006; Kent

et al., 1997; Lowit-Leuschel & Docherty, 2001; Schalling & Hartelius, 2004).

Conceivably, being able to quantify rhythmic differences in an objective and reliable way would facilitate the differential diagnosis of the dysarthrias and perhaps provide a way to track speech progress with recovery and treatment. Furthermore, it is likely that not all patterns of rhythm disturbance are equally detrimental to intelligibility; rather, some rhythmic disorders may prove more problematic to the listener than others to the extent that they impede the successful application of cognitive-perceptual strategies for speech processing, such as the utilization of cues to segmentation of speech into discrete words (Liss, Spitzer, Caviness, Adler, & Edwards, 1998, 2000; Liss, Spitzer, Caviness, & Adler, 2002). This is important to establish, both for advancing our general understanding of the role of rhythm in speech perception and for determining which patterns of rhythm disturbance should be a target for remediation in clinical practice.

A number of duration-based metrics designed to capture differences in speech rhythm between and within languages have recently been developed (Dellwo, 2006; Low, Grabe, & Nolan, 2000; Ramus, Nespor, & Mehler, 1999), all derived from an acoustically based segmentation of the speech signal into vocalic and consonantal (i.e., intervocalic) intervals. No account is taken of syllabic or higher-level prosodic structure, so that all the consonants between successive vowels are included within the same consonantal interval.

Ramus et al. (1999) derived standard deviations of vocalic and consonantal interval durations (ΔV and ΔC , respectively) and found that these metrics, together with %V—the proportion of total utterance duration made up of vowels rather than consonants—were effective in differentiating languages such as Dutch and English from languages such as French and Spanish that had previously been held to be rhythmically distinct.

Several studies have shown that scores for the standard deviation metrics (ΔV , ΔC) are inversely proportional to speech rate (Barry, Andreeva, Russo, Dimitrova, & Kostadinova, 2003; Dellwo & Wagner, 2003; White & Mattys, 2007a), which makes comparison between speakers problematic. Ramus (2002) proposed a simple rate normalization procedure, dividing the standard deviation of interval duration by the mean. This was implemented for consonantal intervals by Dellwo (2006) and was applied to both vocalic and consonantal intervals by White and Mattys (2007a), who found that normalized standard deviation metrics (VarcoV for vowels, VarcoC for consonants) were indeed robust to variation in speech rate.

Developed in parallel with standard deviation measures, pairwise variability indices (PVIs) also derive from vocalic and consonantal interval durations. Here, however,

there is an attempt to capture the syntagmatic nature of rhythm by summing differences between successive intervals on the basis that a high temporal stress contrast language such as English will tend to have much greater durational differences between successive syllables than a language such as Spanish. As for VarcoV, the PVI for vowels is rate-normalized (nPVI-V), whereas the PVI for consonants is not (rPVI-C; see, e.g., Grabe & Low, 2002, for a full account of these metrics).

White and Mattys (2007a, 2007b) compared all of these metrics and found that VarcoV and %V were (a) the most discriminative between languages and between varieties of English held to be rhythmically distinct, (b) the most informative about the influence of first language on second language rhythm, and (c) the most robust to variation in speech rate. Scores for VarcoV and nPVI-V, both rate-normalized metrics of vocalic interval variation, were highly correlated, but the former (a globally normalized metric) was somewhat more discriminative than the latter (a locally normalized metric).

All of the above metrics are based on a division of speech into vocalic and consonantal intervals, from which separate measures of variation are derived. The primary rationale for this is that phonotactic constraints on consonant clustering may vary between languages independently of the degree of vowel reduction/shortening. Thus, for example, Catalan, like Spanish, has minimal consonant clusters but, like English, Catalan significantly reduces and shortens unstressed vowels; in contrast, Polish has complex consonant clusters but little vowel reduction.

This phonological motivation for distinguishing vocalic and consonantal intervals does not apply to dysarthric studies, where rhythmic production is affected instead by articulatory constraints. Furthermore, as discussed previously, the perceptual experience of dysarthric rhythm is generally described as relating to the production of successive syllables. Therefore, in this study we also employed three new rhythm metrics, all based on the duration of successive combined vocalic and consonantal intervals, as an approximation to syllable duration: (a) VarcoVC, the normalized standard deviation; (b) nPVI-VC, the normalized pairwise variability index; and (c) rPVI-VC, the raw (i.e., non-rate-normalized) PVI. Some previous studies have utilized different composite metrics (e.g., Barry et al., 2003; Gut, 2003). We chose not to use pure syllable durations, as in Gut (2003), in order to preserve the acoustic nature of the metrics and to avoid imposing phonological constraints such as syllabification rules on the speech string. We combined each vocalic interval with the subsequent consonantal interval (rather than vice versa, as in Barry et al., 2003), on the basis that phonological theory and psychological research indicate that it is the nature of the syllable rhyme (the vocalic

nucleus plus consonantal coda) that gives rise to the perception of syllable weight.

In summary, we utilized both metrics based on separate vocalic and consonantal intervals and new metrics based on the combined vowel + consonant interval to determine the extent to which they are useful in distinguishing dysarthric speech from that produced by neurologically healthy control speakers as well as distinguishing among four dysarthria subtypes: ataxic, hypokinetic, hyperkinetic, and mixed flaccid-spastic. Specifically, will rhythm metrics capable of distinguishing languages with high and low temporal stress contrast also distinguish among dysarthric speakers of English with perceptually distinct rhythm patterns, and if so, which metrics best succeed in classifying speakers into their respective categories?

Method Speakers

Fifty-five speakers selected from a pool for a larger study provided speech samples that were analyzed for the current study: Twelve with a diagnosis of ataxic dysarthria secondary to various neurodegenerative diseases (Ataxic), 9 with hypokinetic dysarthria secondary to idiopathic PD, 12 with hyperkinetic dysarthria secondary to Huntington's disease (HD), 10 with a mixed spastic-flaccid dysarthria secondary to amyotrophic lateral sclerosis (ALS), and 12 neurologically healthy speakers (Control).¹ Speaker age, gender, and speech descriptions are provided in Table 1. The speakers with dysarthria were selected because their speech deficits were of at least moderate severity (as per intelligibility measures conducted for the larger investigation) and because their perceived symptoms coincided with the cardinal speech features associated with the corresponding speech diagnosis (see Table 1). The presence of the cardinal speech features was established as part of the research protocol, which involved independent perceptual assessment by at least two certified speech-language pathologists.

Speech Stimuli

All speech stimuli were recorded as part of the larger investigation and were obtained within one session (on a speaker-by-speaker basis). Participants were fitted with a head-mounted microphone (Plantronics DSP-100), seated in a sound-attenuating booth, and read stimuli from visual prompts on a computer screen. Recordings were made using a custom script in TF32 (Milenkovic,

¹For convenience, the group codes (PD, HD, ALS, and Ataxic) are based on the associated disease process. It should be recognized, however, that the group membership is based foremost on the speech diagnosis (hypokinetic, hyperkinetic, mixed flaccid-spastic, or ataxic dysarthria).

Table 1. Speech features by dysarthria type, gender of speaker, and age breakdown.

Speaker group	Cardinal perceptual symptoms present, to varying degrees, in all speakers with dysarthria	Gender	Age range M
ALS (N = 10)	Prolonged syllables; slow articulation rate, imprecise articulation; hypernasality; strained strangled vocal quality.	F = 6 M = 4	46–86 X = 64
Ataxic (N = 12)	“Scanning” speech; imprecise articulation with irregular articulatory breakdown; irregular pitch and loudness changes.	F = 6 M = 6	46–87 X = 65
PD (N = 8)	Rapid articulation rate; rushes of speech; imprecise articulation; monopitch; reduced loudness; breathy voice.	F = 2 M = 6	54–81 X = 68
HD (N = 12)	Irregular pitch and loudness changes; irregular rate changes across syllable strings.	F = 6 M = 6	37–80 X = 55
Control (N = 12)	n/a	F = 6 M = 6	21–65 X = 33

Note. ALS = amyotrophic lateral sclerosis; F = female; M = male; X = mean age; PD = Parkinson’s disease; HD = Huntington’s disease; n/a = not applicable.

2004; 16-bit, 44 kHz) and were saved directly to disc for subsequent editing using commercially available software (SoundForge; Sony Corporation, Palo Alto, CA) to remove any noise or extraneous articulations before or after the target utterances. The speakers read 80 short phrases followed by five full sentences. Speakers were encouraged to produce the speech in their “normal, conversational voice.”

The five sentences (Set 1) were adapted by White and Mattys (2007a) from a larger set (Nazzi, Bertoni, & Mehler, 1998). The use of these sentences enabled the direct comparison of data from the present study with previously published work on English (White & Mattys, 2007a, 2007b). The sentences, consisting of 9–12 words (15–17 syllables), were constructed to exclude the approximants /l/, /r/, /w/, and /j/ due to the difficulty of distinguishing these segments from preceding or following vowels on visual waveforms and spectrograms displays. The full set of sentences is listed in Appendix A.

As a source of comparison to the Set 1 sentences, we also collected rhythm metrics on the 80 phrases (Set 2), which were developed for the larger study of intelligibility and lexical boundary errors in dysarthric speech (see Appendix A). In contrast to Set 1 sentences, these phrases were designed to be produced with some rhythmic regularity, either trochaic or iambic stress patterns. The phrases all contained six syllables and were composed of three to five mono- and disyllabic words (based on Liss et al., 1998). The phrases alternated strong (S) and weak (W) syllables (40 each, SW and WS), where *strong syllables* were defined as those carrying lexical stress in citation form. The occurrence of approximants was not controlled in this set of phrases; however, as seen in the paragraphs that follow, the reliability of the derived metrics was highly acceptable. Thus, Set 2 served as a validation set for findings derived from the Set 1 sentences that were developed for previous studies of speech rhythm.

Further, not all of the speakers provided both Set 1 and Set 2 speech material, as the present study was initiated after the start of the larger investigation. Set 1 included measures of the five sentences produced by 34 dysarthric and 9 control speakers. Set 2 measurements were conducted on four sets of 80 phrases, one for each dysarthria type consisting of a subset of phrases from each of the 40 dysarthric speakers² as well as on a comparable set of phrases produced by 5 control speakers. All participants in the Ataxic and ALS groups provided both sets of speech material, but the participant overlap between Sets 1 and 2 was 89% for the PD group, 67% for the HD group, and 42% for the control group. Thus, in addition to being much shorter and having different metrical structure, Set 2 also provided an opportunity to evaluate Set 1 results on a somewhat different constellation of speakers.

Temporal Measurements

All speech samples were analyzed using Praat (Boersma & Weenik, 2006) and TF32 (Milenkovic, 2004) software. For calculation of the rhythm metrics, CV and VC boundaries were identified and labeled by the fourth and sixth authors (KL and SS, respectively) by visual inspection of speech waveforms and spectrograms according to standard segmentation criteria (Peterson & Lehiste, 1960), with labels placed at the point of zero crossing on the waveform. The primary indicator of a VC boundary was the end of a pitch period preceding a break in the formant structure, with a corresponding drop in waveform amplitude. Vowel offset boundary labels were further placed based on a change in the shape of the successive

²For the purpose of the larger investigation, we selected a subset of 6–8 phrases from each speaker to construct a full set of 80 for each dysarthria group. Phrases were selected based on the presence of cardinal features and a moderate-to-severe intelligibility deficit level.

pitch periods, the onset of visible frication (fricatives), and the onset of nasal formant structure combined with waveform amplitude minima (nasals). CV boundaries were primarily determined as beginning at the start of the pitch period coinciding with the onset of regular formant structure, thus aspiration following stop consonant release was included with the consonant interval; the amplitude, shape, and lack of frication of successive pitch periods was also a guide. Following White and Mattys (2007a), vocalic intervals were identified and measured only when there was visible evidence of a voiced vowel: Devoiced vowels and syllabified consonants were included in the adjacent consonant interval.

Vocalic and consonant interval durations were extracted using a custom Praat script on the boundary label files. Some dysarthric speakers, particularly those in the ALS group, produced numerous pauses for inhalation on the longer sentences of Set 1. Following previous procedure (see White & Mattys, 2007a), these silent pauses were excluded. The durations of successive vowels or consonants were summed to form one interval duration, both when immediately adjacent and when separated by a pause. This standard procedure maintained equality of the numbers of vocalic and consonantal intervals while removing the need for linguistic judgments about prosodic constituency and pre-final lengthening in the calculation of what are intended to be fundamentally acoustic metrics (see Grabe & Low, 2002; Ramus et al., 1999). Phrase-initial consonants, where present, were excluded from the analysis to maintain consistency across phrases.

Reliability

To determine intra- and inter-rater reliability of the rhythm metrics, all Set 1 and Set 2 speech tokens produced by two randomly chosen speakers from each speaker group were re-measured by same and different judges using Praat. The vocalic and consonantal intervals were re-labeled, and the durations of each were compared, using Cronbach's alpha, to the originally obtained metrics. Intra- and inter-rater reliability results are presented in Table 2 for both sets of speech material and for all five speaker groups. For intra-rater reliability, Cronbach's alpha scores ranged from .956 to .986; inter-rater scores ranged from .915 to .991. Both sets of scores were deemed acceptable (they were not systematically poorer for any given dysarthric group, either).

Metrics

Table 3 contains a list of all of the variables entered into the statistical analyses. The first 10 variables in this table are the rhythm metrics, as described in the

Table 2. Interval measurement intra- and inter-rater reliability.

Dysarthria type	Set 1		Set 2	
	Intra-rater (α)	Inter-rater (α)	Intra-rater (α)	Inter-rater (α)
Ataxic	.980	.977	.977	.969
ALS	.979	.982	.982	.981
HD	.986	.964	.972	.972
PD	.976	.950	.977	.991
Control	.968	.982	.956	.915

introduction. In addition, articulation rate was calculated as syllables per second, based on the actual time participants took to utter the syllables of the target phrase or sentence, excluding pauses and dysfluencies.

Analysis

The primary goal of data analysis was to identify whether metrics of speech rhythm could robustly distinguish speakers with dysarthria from healthy control speakers and, further, the extent to which these metrics could distinguish among the different forms of dysarthria. Toward this end, a series of stepwise discriminant function analyses (DFAs) was undertaken using SPSS (Version 15.0). DFA is an ideal tool for the present purpose because it is known to be effective in determining which set of continuous variables (e.g., rhythm metrics) best discriminate between naturally occurring groups (e.g., dysarthrias), providing a quantitative composite index of group membership for each observation (e.g., speakers). At each stage of the stepwise (forward) DFA, the variable that minimized Wilks's lambda was entered into the DFA, provided its F statistic was significant ($p < .05$). At any point during the analysis, variables were removed from the DFA if they were found to be no longer significant ($p > .10$) when a new variable was added.

Canonical functions, representing linear combinations of the selected (i.e., most powerful) predictor variables, were constructed by the DFA and were used to create classification rules for group membership. The accuracy with which these rules classify the members of the group is expressed as a percentage. Because the classification rules are, in essence, tailored to the specific data set, it is necessary to invoke a more stringent test to assess reliability of the original classification results. Here, we employed cross-validation (also called the "leave-one-out method"). By this method, the DFA constructs the classification rules using all but one of the speakers. The excluded speaker is then classified based on the functions derived from all other speakers. This is repeated for all speakers, and the resulting classification accuracy, which is usually lower than the original classification

Table 3. Definitions of rhythm metrics and articulation rate measure derived for all speech materials.

Measure	Description
ΔV	Standard deviation of vocalic intervals.
ΔC	Standard deviation of consonantal intervals.
%V	Percent of utterance duration composed of vocalic intervals.
VarcoV	Standard deviation of vocalic intervals divided by mean vocalic duration ($\times 100$).
VarcoC	Standard deviation of consonantal intervals divided by mean consonantal duration ($\times 100$).
VarcoVC	Standard deviation of vocalic + consonantal intervals divided by mean vocalic + consonantal duration ($\times 100$).
nPVI-V	Normalized pairwise variability index for vocalic intervals. Mean of the differences between successive vocalic intervals divided by their sum ($\times 100$).
rPVI-C	Pairwise variability index for consonantal intervals. Mean of the differences between successive consonantal intervals.
nPVI-VC	Normalized pairwise variability index for vocalic + consonantal intervals. Mean of the differences between successive vocalic + consonantal intervals divided by their sum ($\times 100$).
rPVI-VC	Pairwise variability index for vocalic and consonantal intervals. Mean of the differences between successive vocalic and consonantal intervals.
Articulation rate	Number of (orthographic) syllables produced per second, excluding pauses.

Note. For full details of pairwise variability index (PVI) calculations, see Grabe and Low (2002).

accuracy, provides an index of robustness of the original DFA results.

Three analyses were conducted on Set 1, and the results from each were applied to Set 2 for verification (in a nonstepwise DFA). Analysis 1 focused on seven previously used rhythm metrics: ΔV , ΔC , %V, VarcoV, VarcoC, nPVI-V, and rPVI-C. In Analysis 2, the newly defined composite metrics—VarcoVC, nPVI-VC, and rPVI-VC—were included with the seven “standard” metrics used in Analysis 1. Finally, Analysis 3 included all of the previous metrics plus articulation rate. Because articulation rate was, as expected, highly correlated with many rhythm metrics, it was not included in the first two analyses to achieve an independent assessment of the other predictor variables. Stepwise DFA is sensitive to multicollinearity, so the analysis preferentially selects only one of the highly correlated metrics as a variable for the classification function³ (correlation coefficients are provided in Appendixes B and C for reference).

Results

One-Way Analyses of Variance (ANOVAs)

To determine if the metrics demonstrated significant group differences, a series of one-way ANOVAs was conducted. All of the metrics established significant group differences ($p < .01$). To ensure that the group differences

were not simply due to a broad contrast with the control speakers, another set of ANOVAs was completed for the dysarthric speaker groups only. All metrics were significantly different across groups ($p < .05$; see Table 4; see also Appendix D for eigenvalues). On the basis of these results, all variables were considered in their respective stepwise DFAs. Mean values and standard error of measurement for each variable per speaker group are reported in Table 5.

Stepwise Discriminant Function Analyses: Overall Data

Analysis 1

The following metrics were input into the stepwise DFA for Set 1: ΔV , ΔC , %V, VarcoV, VarcoC, nPVI-V, and rPVI-C. This DFA identified four of these seven variables as being most important for maximizing the distances among group distributions, and these were entered into the classification function as follows, in order of the most to least important: VarcoV, ΔV , VarcoC, and %V. A total of four canonical functions, which represent linear combinations of the entered variables, were constructed. The classification function derived from the DFA accurately classified 79% of the speakers into their respective speaker groups (74% with the cross-validation method).

The four predictor variables identified for Set 1 served as input variables for the (nonstepwise) DFA on Set 2 (VarcoV, ΔV , VarcoC, and %V). The classification rules based on these variables resulted in correct classification of 82% of the speakers (73% with the cross-validation method; see Table 6 for summary of correct and incorrect classifications for Analysis 1).

Thus, the results of this first DFA demonstrate the ability of established rhythm metrics to distinguish among speaker groups. The predictor variables identified in the

³Note that because of DFA's sensitivity to collinearity, the set of selected variables is not necessarily the only set that will result in high classification accuracy.

Table 4. Results of one-way analyses of variance (ANOVAs) testing equality of means for set 1 and set 2, and for all speakers and dysarthria-only.

Variable	Set 1: All speakers		Dysarthria only		Set 2: All speakers		Dysarthria only	
	F(4, 38)	p	F(3, 30)	p	F(4, 40)	p	F(3, 35)	p
ΔV	15.654	.000	9.593	.000	5.530	.001	5.462	.003
ΔC	14.562	.000	10.332	.000	10.526	.000	10.101	.000
% V	5.975	.001	3.183	.038	6.708	.001	4.210	.012
nPVI-V	12.197	.000	5.903	.003	9.546	.000	6.442	.001
rPVI-C	15.240	.000	11.230	.000	7.850	.000	7.375	.001
VarcoV	17.868	.000	10.067	.000	9.231	.000	6.687	.001
VarcoC	10.092	.000	10.884	.000	7.035	.000	7.718	.000
rPVI-VC	16.046	.000	10.942	.000	7.196	.000	7.277	.001
nPVI-VC	8.574	.000	9.570	.000	5.829	.000	6.778	.001
VarcoVC	10.471	.000	11.795	.000	7.249	.000	8.736	.000
Artic. rate	81.545	.000	56.123	.000	62.222	.000	76.807	.000

stepwise analysis also were successful for high classification accuracy with Set 2, showing the extent to which these variables are robust on a very different type of speech material and with a slightly different constellation of speakers. Thus, the combination of VarcoV, ΔV , VarcoC, and %V is capable of separating the distributions

among the five speaker groups for this speech material as well as generalizing to a new set. It is of note that this set of metrics includes VarcoV and %V, found by White and Mattys (2007a, 2007b) to be the most discriminative for rhythmically distinct languages and robust to articulation rate variation.

Table 5. Mean values by speaker group per variable (metric).

Variable	Speaker group				
	PD M (SE)	HD M (SE)	ALS M (SE)	Ataxic M (SE)	Control M (SE)
Set 1					
ΔV	48.19 (2.89)	73.55 (11.98)	75.49 (3.73)	73.43 (3.00)	44.57 (1.10)
ΔC	55.33 (3.24)	127.33 (22.27)	108.50 (11.07)	95.82 (4.50)	55.77 (1.26)
%V	46.31 (1.32)	45.83 (2.18)	51.82 (1.53)	45.87 (0.95)	41.48 (0.53)
nPVI-V	59.41 (2.21)	56.68 (3.08)	42.57 (2.09)	54.50 (2.01)	66.57 (1.33)
rPVI-C	63.89 (3.56)	144.77 (26.67)	126.38 (11.80)	109.17 (4.70)	65.42 (1.85)
VarcoV	50.49 (1.31)	48.85 (4.23)	34.85 (1.48)	44.11 (1.55)	55.59 (0.98)
VarcoC	49.75 (1.95)	68.11 (4.86)	51.09 (2.55)	48.38 (1.85)	48.84 (1.38)
rPVI-VC	78.48 (4.78)	172.51 (32.03)	141.80 (12.29)	132.59 (6.10)	74.27 (2.53)
nPVI-VC	37.98 (1.35)	47.53 (3.32)	32.02 (1.67)	35.18 (1.17)	38.00 (0.97)
VarcoVC	32.42 (2.76)	45.42 (3.27)	28.94 (1.50)	32.06 (1.28)	33.14 (0.93)
Artic. rate	4.98 (0.20)	3.28 (0.36)	2.45 (0.16)	2.80 (0.11)	5.20 (0.11)
Set 2					
ΔV	43.88 (2.76)	78.62 (8.69)	80.86 (5.85)	77.47 (6.33)	51.34 (3.96)
ΔC	53.26 (3.94)	104.45 (8.25)	100.65 (7.35)	92.36 (5.8)	61.12 (2.0)
%V	47.62 (1.81)	46.77 (1.77)	53.07 (1.84)	45.59 (1.27)	38.41 (1.65)
nPVI-V	58.12 (2.42)	57.42 (4.07)	39.27 (2.01)	54.87 (3.89)	76.08 (4.56)
rPVI-C	62.13 (4.34)	115.72 (9.53)	115.52 (9.68)	102.54 (7.68)	68.51 (3.11)
VarcoV	44.40 (1.81)	46.14 (3.05)	31.75 (1.33)	40.30 (1.76)	55.50 (3.13)
VarcoC	47.44 (1.35)	52.63 (2.43)	44.88 (1.73)	40.34 (1.3)	40.98 (1.85)
rPVI-VC	81.93 (4.97)	160.25 (15.3)	147.13 (11.3)	152.50 (11.38)	100.78 (5.2)
nPVI-VC	39.47 (2.16)	42.56 (2.47)	30.40 (1.09)	36.23 (1.34)	41.70 (2.09)
VarcoVC	33.57 (1.38)	37.74 (2.07)	26.03 (.9)	30.14 (1.4)	34.52 (1.34)
Artic. rate	4.85 (.15)	2.93 (0.11)	2.16 (0.13)	2.48 (0.11)	4.23 (.25)

Table 6. Classification summary from analysis 1.

Speaker group	Predicted group membership					Total
	Ataxic	ALS	HD	PD	Control	
Set 1						
<i>Original count</i>						
Ataxic	11	1				12
ALS	2	7	1			10
HD			4			4
PD				5	3	8
Control				2	7	9
<i>Cross-validated count</i>						
Ataxic	10	2				12
ALS	2	7	1			10
HD		3	1			4
PD				5	3	8
Control				2	7	9
Set 2						
<i>Original count</i>						
Ataxic	11				1	12
ALS	1	7				8
HD	2		7	1	2	12
PD				7	1	8
Control					5	5
<i>Cross-validated count</i>						
Ataxic	10	1	1			12
ALS	2	6				8
HD	2		7	1	2	12
PD				7	1	8
Control			1	1	3	5

Note. For Set 1, 79% of originally grouped speakers and 74% of cross-validated speakers correctly classified. For Set 2, 82% of originally grouped speakers and 73% of cross-validated speakers correctly classified.

Analysis 2

VarcoVC, nPVI-VC, and rPVI-VC were added to the rhythm metrics used in Analysis 1. Five of the 10 variables were entered into the stepwise DFA in order of importance: VarcoV, VarcoVC, ΔV , %V, and ΔC . A total of four canonical functions was constructed by the DFA, the first three accounting for 98% of the variance. With addition of the new metrics into the analysis, 88% of speakers were correctly classified. Cross-validation correctly classified 77% of the speakers.

As before, the predictor variables identified in Set 1 were entered into a DFA on Set 2. VarcoV, VarcoVC, ΔV , %V, and ΔC resulted in correct classification of 78% of the speakers. The cross-validation procedure resulted in correct classification of 69% of the speakers (see Table 7 for a detailed summary of classification results).

The results of Analysis 2 suggest that composite metrics may be important for distinguishing among speaker

groups, with Set 1 classification improved relative to Analysis 1. In particular, VarcoVC emerged as a reliable predictor variable. The Set 2 results demonstrate the generalizability of this set of predictor variables to new speech materials and speakers.

Analysis 3

The inclusion of articulation rate in this DFA for Set 1 caused the model to drop four of the metrics that were important in Analyses 1 and 2: VarcoV, ΔV , %V, and ΔC . This elimination was the result of the high correlations between these metrics and articulation rate. The variables that were then entered into the DFA included articulation rate, VarcoC, nPVI-V, and VarcoVC. Four canonical functions were created and used to classify the data, the first three of which accounted for 98% of the total variance. Correct classification was obtained for 79% of the originally grouped speakers and 72% of the cross-validated speakers.

Table 7. Classification summary from analysis 2.

Speaker group	Predicted group membership					Total
	Ataxic	ALS	HD	PD	Control	
Set 1						
<i>Original count</i>						
Ataxic	12					12
ALS	2	8				10
HD			3	1		4
PD				6	2	8
Control					9	9
<i>Cross-validated count</i>						
Ataxic	9	2			1	12
ALS	2	7	1			10
HD			3	1		4
PD				5	3	8
Control					9	9
Set 2						
<i>Original count</i>						
Ataxic	11		1			12
ALS	2	6				8
HD	2		7	1	2	12
PD				7	1	8
Control				1	4	5
<i>Cross-validated count</i>						
Ataxic	9	2	1			12
ALS	2	6				8
HD	2	1	6	1	2	12
PD				7	1	8
Control	1			1	3	5

Note. For Set 1, 88% of originally grouped speakers and 77% of cross-validated speakers correctly classified. For Set 2, 78% of originally grouped speakers and 69% of cross-validated speakers correctly classified.

A DFA was conducted on Set 2, with the predictor variables identified with Set 1 as input. Using articulation rate, VarcoC, nPVI-V, and VarcoVC, the classification rules correctly classified 89% of the originally grouped speakers and 82% of the cross-validated speakers (see Table 8 for a summary of classification results).

In summary, classification results for both Set 1 and Set 2 were of generally high accuracy for all analyses. For Set 1, the best classification was obtained with the group of variables in Analysis 2: VarcoV, VarcoVC, ΔV , %V, and ΔC ; for Set 2, the best classification derived from the Analysis 3 group, which included articulation rate. Together, these analyses confirm the ability of rhythm metrics to yield distinguishable distributions among the speaker groups in the DFA.

Misclassifications

Misclassifications (see Tables 6–8) can be summarized as follows: Control speakers were rarely misclassified, and

only as PD; PD speakers were only misclassified as control; Ataxic speakers were misclassified as ALS (with a single HD classification); ALS speakers were misclassified only as ataxic; and HD speakers were misclassified as Ataxic, ALS, and control. To determine which predictor variables are most useful for distinguishing one dysarthric speaker group from the rest, four additional DFAs were conducted.

Stepwise DFAs: Dysarthria-Specific Comparisons

Analysis 4: Hypokinetic Dysarthria (PD)

In this analysis, PD was pitted against the other dysarthric groups combined. All 10 rhythm metrics and articulation rate were input to the DFA. The stepwise analysis entered articulation rate, nPVI-V, and ΔV as predictor variables whose first canonical function resulted in 100% classification accuracy for Set 1. Given this finding, we conducted two subsequent nonstepwise DFAs on both

Table 8. Classification summary from analysis 3.

Speaker group	Predicted group membership					Total
	Ataxic	ALS	HD	PD	Control	
Set 1						
<i>Original count</i>						
Ataxic	11	1				12
ALS	3	6	1			10
HD			4			4
PD				5	3	8
Control				1	8	9
<i>Cross-validated count</i>						
Ataxic	10	2				12
ALS	3	6	1			10
HD		1	3			4
PD				5	3	8
Control				2	7	9
Set 2						
<i>Original count</i>						
Ataxic	11	1				12
ALS		8				8
HD	2	1	9			12
PD				7	1	8
Control					5	5
<i>Cross-validated count</i>						
Ataxic	11	1				12
ALS	1	7				8
HD	2	1	9			12
PD				7	1	8
Control				2	3	5

Note. For Set 1, 79% of originally grouped speakers and 72% of cross-validated speakers correctly classified. For Set 2, 89% of originally grouped speakers and 82% of cross-validated speakers correctly classified.

Sets 1 and 2, with articulation rate as the only input variable. These also achieved 100% classification accuracy in distinguishing PD from all other dysarthric speakers.

Interpretation. These results are consistent with the perceptual characterization of the PD speech as rapid or rushed, and this was supported by the articulation rate values, whose means were similar to or exceeded the control rates (see Table 5). It is believed that the reduced excursions of the articulators secondary to basal ganglia dysfunction in PD give rise to the perceptual experience of rushed speech in the context of normal or supranormal rates (Caligiuri, 1989; Weismer, 1984). It is of particular interest that despite the reduced intelligibility, the metrics that captured temporal relationships among vocalic and consonantal segments—in particular, VarcoVC—showed relative preservation of normal rhythm, as evidenced by the similarity of scores to those of the control group. This is consistent with observations in other motor systems in PD, which show that the motor program for a

task is essentially intact but that it is implemented in a scaled-down spatial domain (Berardelli, Dick, Rothwell, Day, & Marsden, 1986; Hallett & Khoshbin, 1980).

Analysis 5: Ataxic Dysarthria (Ataxic)

Two categories of speakers were established: Ataxic versus all other dysarthric speakers. The stepwise DFA using all metrics plus articulation rate identified VarcoC, rPVI-VC, and nPVI-V as predictor variables. The function derived from the first, and only, linear combination of these variables resulted in 85% accurate classifications, with 79% accuracy on cross-validation. A separate stepwise DFA conducted on Set 2 identified VarcoC but did so also along with ΔC and %V as predictor variables. Application of this function resulted in 85% accurate classification for both the original and cross-validated methods. The ranking of the predictor variables in both cases suggests that the rate-normalized consonantal interval measure (VarcoC) was useful for distinguishing ataxic

speech from the other dysarthrias. The explanation for this lies in Table 5, which shows that the mean VarcoC values for this group were highly similar to those for the control group (not included in this analysis). Thus, the relative normality of this metric distinguished ataxia from the other dysarthrias, much in the same way that the relative normality of articulation rate distinguished PD from the rest of the dysarthric groups.

Interpretation. Ataxia arising from disruptions in cerebellar circuitry is associated with disturbances in rhythm and movement timing and coordination. In the speech domain, the ataxia manifests as irregularities in production as well as in a tendency toward equal and even scanning speech rhythm. Metrics previously held to capture the scanning quality of ataxic connected speech (such as nPVI-V or nPVI-VC) did not emerge in this analysis as the primary variables for distinguishing ataxia from the other dysarthrias (nPVI-V was the third predictor variable for Set 1 only). These “scanning” metrics were indeed low as compared with control (see Table 5), and the absolute values are highly similar to those reported in previous studies that document scanning in ataxia (Henrich et al., 2006). But these metrics did not emerge as powerful variables because the values for the ataxic speech were not as low as those derived from the ALS speech, which exhibited more severe metrical scanning. Also of note in Table 5 (and referencing Appendix B) is that the mean values for ΔC , ΔV , and rPVI-VC (all strongly positively correlated) were high relative to control but were similar to the mean values for the HD and ALS groups. Thus, at least in distinguishing these speakers with ataxic dysarthria from those with other dysarthria subtypes, the patterns of abnormality were not distinctive. However, with the canonical functions derived from a linear combination of predictor variables, classification accuracy for ataxia was very good, particularly for Analysis 3, in which only 1 ataxic speaker was misclassified on cross-validation. This highlights the value of an analysis that is able to combine multiple predictor variables to attain an effective solution.

Analysis 6: Flaccid-Spastic Dysarthria (ALS)

Two categories of speakers were established: ALS versus all other dysarthric speakers. The stepwise DFA using all metrics plus articulation rate identified only VarcoV as a predictor variable. The classification function was 82% accurate in classification, both with the original and cross-validated groupings. The stepwise DFA conducted on Set 2 identified metrics nPVI-V (highly correlated with VarcoV; refer to Appendix B) and %V. The function achieved 85% classification accuracy for both the original and cross-validated groupings. It is remarkable that the high levels of original and cross-validated

classification accuracy were identical within both Set 1 and Set 2, indicating the robustness of these correlated variables in classification. These analyses converge in suggesting that metrics which capture temporal variation of vocalic segments are useful in distinguishing ALS from the other dysarthric groups. Indeed, rhythm scores for ALS showed the lowest degree of vocalic interval variation, as evidenced by the low scores for rate-normalized metrics (see Table 5; VarcoV, nPVI-V, and also VarcoVC).

Interpretation. The mixed flaccid-spastic dysarthria of ALS is the result of both lower and upper motor neuron degeneration. Speech is slow and prolonged, and breath groups are small due to deficits in respiratory support and the valving challenges caused by glottal stenosis (strained-strangled vocal quality) and impaired velopharyngeal function. Movement velocities of articulators, especially the tongue, are slow (Weismer, Yunusova, & Westbury, 2003; Yunusova, Weismer, Westbury, & Lindstrom, 2008). These features were borne out in the rhythm data. The most important predictor variables for ALS were those that captured the prolongation of vowels (%V) and the lack of temporal distinction between vowels produced in stressed versus unstressed syllables (nPVI-V, VarcoV). Thus, the ALS speech was the most syllable-by-syllable of the four dysarthria groups.

Analysis 7: Hyperkinetic Dysarthria (HD)

Two categories of speakers were established: HD versus all other dysarthric speakers. The stepwise DFA using all metrics plus articulation rate identified VarcoC as the only predictor variable. The classification function was 85% accurate in classification for both the original and cross-validated groupings. The stepwise DFA conducted on Set 2 identified VarcoVC (significantly correlated with VarcoC; see Appendix B) as the only predictor variable. Classification based on the first function was 75% accurate for both original and cross-validated groupings. With the highest VarcoC and VarcoVC values of any group, these metrics captured the high variability of consonantal segment durations. The same variability was not as pronounced in the vocalic intervals. Other metrics that captured the consonantal variability that did not emerge in the analysis but that are represented with high values included ΔC , rPVI-C, rPVI-VC, and nPVI-VC. As for ataxia and ALS, articulation rate was not discriminative for HD: All of these dysarthrias have reduced rates of speech.

Interpretation. Darley and colleagues (1969) included the perceptual symptom of “bizarreness” to characterize the highly unusual sound of hypokinetic dysarthria of the choreic type. The seemingly random bursts of vocalization and intrusive and extraneous orofacial movements result in variable breakdown of the forward flow of speech

(Hartelius, Carlstedt, Ytterberg, Lillvik, & Laakso, 2003). In the present study, metrics sensitive to the high variability in the consonantal intervals (VarcoC and VarcoVC) were important for distinguishing HD speech from that of the other dysarthrias. Other variables that also captured this variability included ΔC , rPVI-C, rPVI-VC, and nPVI-C. It is also of note that the longer speech material of Set 1 was more useful than the Set 2 short phrases in distinguishing HD from the other dysarthrias. Because of the intermittent manifestation of the choreic movements, we can speculate that the longer material offered more opportunity to capture the effects of the hyperkinesia.

Discussion

The first question posed by this investigation was whether rhythm metrics could be successful at distinguishing the speech of healthy control participants from that of persons with dysarthria. Overall, there was good success at this, with nearly 80% correct classification of control speakers when we collapse across all of the first three sets of analyses. A particularly impressive result was that returned for Analysis 2 (Set 1), in which there was 100% classification accuracy for control speakers, even on cross-validation. The success of this analysis may be linked to the inclusion of the new rhythm metrics, which were intended to be sensitive to disruptions in the forward flow of dysarthric speech (VarcoVC was the second strongest predictor variable in this set). Second, and most interestingly, we asked whether predictor variables achieved accurate differential classification among the dysarthrias. The majority of classification functions reported herein were more than 80% successful in classifying speakers into their appropriate group; and the more stringent cross-validation methods were more than 70% successful. Further insight is provided by the clusters of variables that emerged when pitting each dysarthria against the rest (Analyses 4–7). These variables coincide with perceptual features and underlying production constraints associated with the diagnostic categories (Kent & Kim, 2003). Thus, to best classify an entire group of speakers that includes control and the dysarthrias, metrics sensitive to the particular patterns of rhythm generated by each of the groups should be included.

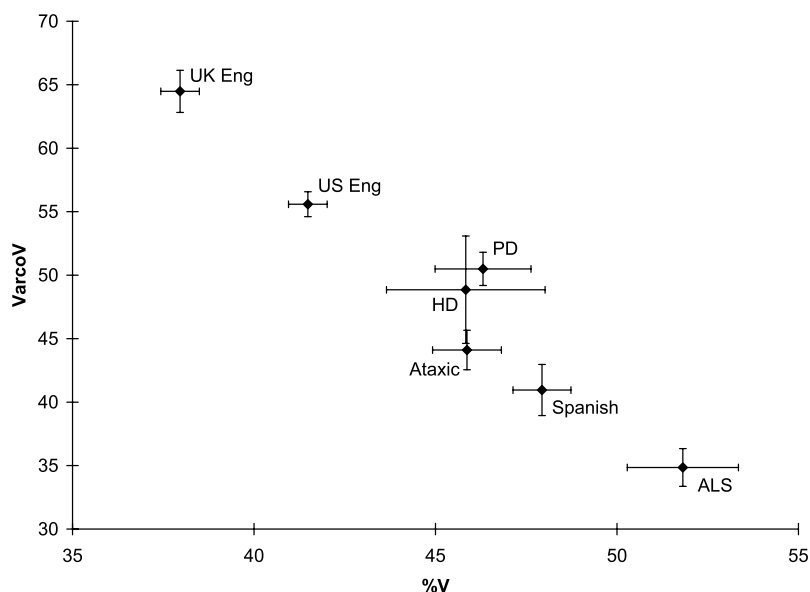
This investigation has several clinical implications. First, the overall success in classification suggests that rhythm metrics may provide an objective means to augment differential diagnosis. This is an exciting possibility, but more studies are needed to determine whether mild presentations are distinguishable from healthy speech and, further, whether mild (or very severe) presentations lend themselves to dysarthria subtype classification. The present study can be regarded as a rather optimal data

set for classification because the speech was moderate to severely affected, and members of each group presented with at least some level of the cardinal perceptual features. It is expected that, similar to perceptual evaluation, the application of rhythm metrics to differential diagnosis will have both limitations and strengths. An important line of investigation will be to determine whether rhythm metrics can exceed the sensitivity and specificity of auditory perceptual judgments across the full range of presentation severity.

A second and perhaps more immediately useful function of rhythm metrics is their capacity to objectively track speech change over the course of an individual's disease progression, or improvements associated with pharmacological, surgical, or behavioral interventions. As the present study shows, the acquisition of rhythm metrics is fairly straightforward, and only a small corpus of speech material (e.g., five sentences) is needed. Measurements of vocalic and consonantal intervals can be operationally defined and reliably measured by those familiar with acoustic analysis and spectrographic displays of speech. The rhythm metrics then can be easily calculated using any spreadsheet software. The efficacy of rhythm metrics for clinical use would be further increased if the measurement process could reliably be automated. Methods of automated alignment of the segment boundaries of transcribed speech are available and, with further improvement and appropriate modifications to accommodate disordered speech, may lead to a rational, objective, and effective clinical tool.

Third, an important practical and theoretical question that emerges from these data concerns the ways in which these different patterns of rhythmic disturbance contribute to associated decrements in intelligibility. In particular, to what extent do intelligibility deficits reflect a listener's inability to cope with degraded or misleading rhythm cues, causing them to make errors in identifying word boundaries? Although a full treatment of this issue is beyond the scope of this article, the co-investigation of lexical boundary errors and rhythm metrics may provide useful insights into a source of intelligibility decrement in the dysarthrias. It may also point toward targets for rhythm remediation by revealing how the various patterns differentially contribute to lexical boundary errors. For example, we can speculate that rhythmic disturbances that obscure durational differences between stressed and unstressed syllables will have the most deleterious effect on lexical segmentation (at least in English). Because the PD values indicated a relative preservation of timing patterns, we might expect to see relatively less difficulty applying a lexical segmentation strategy (support for this is presented in Liss et al., 2000, 2002). In contrast, the highly variable timing relations in HD, or the extreme scanning rhythm of ALS, would be detrimental to its application because the

Figure 1. VarcoV as a function of %V for Set 1. Means and standard error bars for the present speakers are plotted along with data from previously published work in British English and Catalan Spanish (White & Mattys, 2007a). Note that all dysarthric data fall in the direction of reduced temporal contrast relative to both British and American English, with amyotrophic lateral sclerosis (ALS) speech being the most extreme, even exceeding the values for Spanish rhythm. %V = percent of utterance duration composed of vocalic intervals.



durational differences would be more random (in HD) or reduced (in ALS). In the latter two cases, improvements of temporal relationships, perhaps through contrastive stress training, would be useful to address that particular source of intelligibility decrement.

Finally, the present findings raise an interesting question regarding the consequences of rhythm abnormalities in dysarthria across different language classes. Figure 1 shows the VarcoV data for all participants in the present study plotted as a function of %V, relative to previously published values for British English and Catalan Spanish (White & Mattys, 2007a). The values to the upper left (U.K. British) are associated with a high temporal stress contrast rhythmic pattern (a.k.a., “stress-timed”); values to the lower right represent low temporal stress contrast (a.k.a., “syllable-timed”). It is of note that all of the dysarthric speakers in this investigation fell to the lower right of the control speakers (U.S. English) from this investigation, indicating measurable reductions in temporal stress contrast. Indeed, the speakers with ALS produced even less temporal contrast than Spanish. We have suggested and cited evidence that this reduction in temporal contrast is a source of intelligibility decrement for English listeners who rely on this cue for lexical segmentation (e.g., Liss et al., 2000). But would this be the case to the same extent for listeners less inclined to rely on this cue because of its lack of relevance in their own language, such as Spanish or French? It is conceivable that

the rhythm abnormalities in dysarthria—and perhaps other aspects of speech deficit, as well—cause fundamentally different challenges for listeners across languages. Rhythm metrics provide a potentially fruitful platform for the investigation of cross-language differences in communicative impairment secondary to speech production disorders such as the dysarthrias.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communicative Disorders Grant 5 R01 DC 6859, awarded to J. M. Liss (first author) and Grant F/00 182/BG from the Leverhulme Trust to S. L. Mattys (third author). Gratitude is extended to Katy Kennerley and Yu-kyong Choe for their assistance with data analysis. We are especially appreciative for those who supported and participated in this research, including volunteers from Mayo Clinic Arizona; Phil Hardt and members of the Huntington’s Disease Society of America, Arizona Affiliate; the Arizona Ataxia Support Group; Pamela Mathy; and K. Sivakumar and his patients at the Neuromuscular Research Center in Scottsdale, Arizona.

References

- Ackermann, H., & Hertrich, I. (1994). Speech rate and rhythm in cerebellar dysarthria: An acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica*, 46, 70–78.

- Barry, W. J., Andreeva, B., Russo, M., Dimitrova, S., & Kostadinova, T.** (2003). Do rhythm measures tell us anything about language type? In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetics Sciences* (pp. 2693–2696). Rundle Mall: Causal Productions.
- Berardelli, A., Dick, J. P., Rothwell, J. C., Day, B. L., & Marsden, C. D.** (1986). Scaling of the size of the first agonist EMG burst during rapid wrist movements in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery, & Psychiatry*, *49*, 1273–1279.
- Boersma, P., & Weenik, D.** (2006). Praat [Computer software]. Amsterdam: University of Amsterdam.
- Borzone de Manrique, A. M., & Signorini, A.** (1983). Segmental durations and the rhythm in Spanish. *Journal of Phonetics*, *11*, 117–128.
- Caligiuri, M.** (1989). The influence of speaking rate on articulatory hypokinesia in Parkinsonian dysarthria. *Brain and Language*, *36*, 493–502.
- Darley, F. L., Aronson, A. E., & Brown, J. R.** (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research*, *12*, 246–269.
- Dauer, R. M.** (1983). Stress-timing and syllable-timing re-analyzed. *Journal of Phonetics*, *11*, 51–62.
- Dellwo, V.** (2006). Rhythm and speech rate: A variation coefficient for delta C. In P. Karnowski & I. Szigeti (Eds.), *Language and language processing: Proceedings of the 38th Linguistic Colloquium, Piliscsaba 2003* (pp. 231–241). Frankfurt: Peter Lang.
- Dellwo, V., & Wagner, P.** (2003, August). Relations between language rhythm and speech rate. In M. J. Solé, D. Recasens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetics Sciences* (pp. 471–474). Barcelona: Causal Productions.
- Grabe, E., & Low, E. L.** (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner & C. Gussenhoven (Eds.), *Papers in Laboratory Phonology 7* (pp. 515–546). Berlin, Germany: Mouton de Gruyter.
- Gut, U.** (2003). Prosody in second language speech production: The role of the native language. *Fremdsprachen Lehren und Lernen*, *32*, 133–152.
- Hallett, M., & Khoshbin, S.** (1980). A physiological mechanism of bradykinesia. *Brain*, *103*, 301–314.
- Hartelius, L., Carlstedt, A., Ytterberg, M., Lillvik, M., & Laakso, K.** (2003). Speech disorders in mild and moderate Huntington disease: Results of dysarthria assessments of 19 individuals. *Journal of Medical Speech-Language Pathology*, *11*, 1–14.
- Hartelius, L., Runmarker, B., & Andersen, O.** (2000). Prevalence and characteristics of dysarthria in a multiple-sclerosis incidence cohort: Relation to neurological data. *Folia Phoniatrica et Logopaedica*, *52*, 160–177.
- Henrich, J., Lowit, A., Schalling, E., & Mennen, I.** (2006). Rhythmic disturbance in ataxic dysarthria: A comparison of different measures and speech tasks. *Journal of Medical Speech-Language Pathology (Special Issue: Speech Motor Control Conference Proceedings)*, *14*(4), 291–296.
- Kent, R., Kent, J., Rosenbek, J., Vorperian, H., & Weismer, G.** (1997). A speaking task analysis of the dysarthria in cerebellar disease. *Folia Phoniatrica et Logopaedica*, *49*, 63–82.
- Kent, R. D., & Kim, Y. J.** (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, *17*, 427–445.
- Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C.** (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America*, *112*, 3022–3030.
- Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B.** (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *The Journal of the Acoustical Society of America*, *104*, 2457–2566.
- Liss, J. M., Spitzer, S. M., Caviness, J. N., Adler, C., & Edwards, B.** (2000). Lexical boundary error analysis in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America*, *107*, 3415–3424.
- Lloyd James, A.** (1940). *Speech signals in telephony*. London: Pitman & Sons.
- Low, E. L., Grabe, E., & Nolan, F.** (2000). Quantitative characterizations of speech rhythm: “Syllable-timing” in Singapore English. *Language and Speech*, *43*, 377–401.
- Lowit-Leuschel, A., & Docherty, G. J.** (2001). Prosodic variation across sampling tasks in normal and dysarthric speakers. *Logopedics Phoniatrics Vocology*, *26*, 151–164.
- Milenkovic, P.** (2004). TF32 [Computer software]. Madison, WI: University of Wisconsin–Madison.
- Nazzi, T., Bertoni, J., & Mehler, J.** (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 756–766.
- Peterson, G., & Lehiste, I.** (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America*, *32*, 693–703.
- Pike, K.** (1945). *The intonation of American English*. Ann Arbor, MI: University of Michigan Press.
- Ramus, F.** (2002, April). Acoustic correlates of linguistic rhythm: Perspectives. In B. Bel & I. Marlien (Eds.), *Proceedings of Speech Prosody 2002* (pp. 115–120). Aix-en-Provence, France: Laboratoire Parole et Langage.
- Ramus, F., Nespors, M., & Mehler, J.** (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, *73*, 265–292.
- Schalling, E., & Hartelius, L.** (2004). Acoustic analysis of speech tasks performed by three individuals with spinocerebellar ataxia. *Folia Phoniatrica et Logopaedica*, *56*, 367–380.
- Weismer, G.** (1984). Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination. In M. R. McNeil, J. C. Rosenbek, & A. E. Aronson (Eds.), *The dysarthrias: Physiology, acoustics, perception, management* (pp. 101–130). San Diego, CA: College-Hill Press.
- Weismer, G., Yunusova, Y., & Westbury, J. R.** (2003). Inter-articulator coordination in dysarthria: An x-ray microbeam study. *Journal of Speech, Language, and Hearing Research*, *46*, 1247–1261.
- Wenk, B., & Wiolland, F.** (1982). Is French really syllable-timed? *Journal of Phonetics*, *10*, 193–216.
- White, L., & Mattys, S. L.** (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, *35*, 501–522.

White, L., & Mattys, S. L. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in romance phonology: Current issues in linguistic theory series* (pp. 237–257). Amsterdam: John Benjamins.

Yunusova, Y., Weismer, G., Westbury, J. R., & Lindstrom, M. J. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research*, *51*, 596–611.

Received October 8, 2008

Accepted January 7, 2009

DOI: 10.1044/1092-4388(2009/08-0208)

Contact author: Julie M. Liss, Department of Speech and Hearing Science, Arizona State University, Box 870102, Tempe, AZ 85287-0102. E-mail: julie.liss@asu.edu.

Appendix A (p. 1 of 2). Speech material for Sets 1 and 2.

Set 1

1. The supermarket chain shut down because of poor management.
2. Much more money must be donated to make this department succeed.
3. In this famous coffee shop they serve the best doughnuts in town.
4. The chairman decided to pave over the shopping center garden.
5. The standards committee met this afternoon in an open meeting.

Set 2

1. account for who could knock
2. address her meeting time
3. admit the gear beyond
4. advance but sat appeal
5. afraid beneath demand
6. amend estate approach
7. and spoke behind her sin
8. appear to wait then turn
9. assume to catch control
10. attack became concerned
11. attend the trend success
12. avoid or beat command
13. award his drain away
14. balance clamp and bottle
15. beside a sunken bat
16. bolder ground from justice
17. bush is chosen after
18. butcher in the middle
19. career despite research
20. cheap control in paper
21. commit such used advice
22. confused but roared again
23. connect the beer device
24. constant willing walker
25. cool the jar in private
26. darker painted baskets
27. define respect instead
28. distant leaking basement
29. divide across retreat
30. done with finest handle
31. had eaten junk and train
32. embark or take her sheet
33. for coke a great defeat
34. forget the joke below
35. frame her seed to answer
36. functions aim his acid
37. its harmful note abounds
38. hold a page of fortune
39. increase a grade sedate
40. indeed a tax ascent
41. kick a tad above them
42. listen final station
43. mark a single ladder
44. mate denotes a judgment
45. may the same pursued it
46. measure fame with legal
47. mistake delight for heat
48. mode campaign for budget
49. model sad and local

Appendix A (p. 2 of 2). Speech material for Sets 1 and 2.

50. narrow seated member
 51. her owners arm the phone
 52. pain can follow agents
 53. perceive sustained supplies
 54. pick a chain for action
 55. pooling pill or cattle
 56. push her equal culture
 57. rampant boasting captain
 58. remove and name for stake
 59. resting older earring
 60. rocking modern poster
 61. rode the lamp for teasing
 62. round and bad for carpet
 63. rowing farther matters
 64. seat for locking runners
 65. secure but lease apart
 66. signal breakfast pilot
 67. sinking rather tundra
 68. spackle enter broken
 69. or spent sincere aside
 70. stable wrist and load it
 71. submit his cash report
 72. support with dock and cheer
 73. target keeping season
 74. technique but sent result
 75. thinking for the hearing
 76. to sort but fear inside
 77. transcend almost betrayed
 78. unless escape can learn
 79. unseen machines agree
 80. vital seats with wonder
-

Appendix B. Intercorrelations of the Set 1 measurements ($n = 43$).

Metric	1	2	3	4	5	6	7	8	9	10	11
1. ΔV	—	.734**	.600**	-.450**	.710**	-.471**	.253	-.868**	.803**	-.091	-.018
2. ΔC		—	.219	-.470**	.984**	-.452**	.654**	-.747**	.968**	.266	.345*
3. %V			—	-.629**	.196	-.624**	.076	-.514**	.281	-.337*	-.254
4. nPVI-V				—	-.441**	.923**	-.070	.677**	-.418**	.431**	.294
5. rPVI-C					—	-.437**	.631**	-.738**	.962**	.289	.325*
6. VarcoV						—	.038	.753**	-.412**	.501**	.436**
7. VarcoC							—	-.104	.618**	.726**	.816**
8. Artic. rate								—	-.764**	.278	.214
9. rPVI-VC									—	.334*	.372*
10. nPVI-VC										—	.902**
11. VarcoVC											—

Note. See Table 3 for variable descriptions. Artic. = articulation.

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

Appendix C. Intercorrelations of the Set 2 measurements ($n = 45$).

Metric	1	2	3	4	5	6	7	8	9	10	11
1. ΔV	—	.655**	.554**	-.031	.575**	.058	.158	-.735**	.879**	.179	.225
2. ΔC		—	.147	-.254	.963**	-.122	.565**	-.731**	.868**	.244	.292
3. %V			—	-.484**	.127	-.480**	.096	-.377*	.309*	-.267	-.229
4. nPVI-V				—	-.306*	.900**	-.052	.421**	-.064	.611**	.536**
5. rPVI-C					—	-.173	.564**	-.695**	.800**	.189	.210
6. VarcoV						—	.129	.416**	.033	.741**	.734**
7. VarcoC							—	.039	.346**	.546**	.621**
8. Artic. rate								—	-.740**	.288	.252
9. rPVI-vc									—	.367*	.364*
10. nPVI-vc										—	.903**
11. VarcoVC											—

Note. See Table 3 for variable descriptions.

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

Appendix D. Eigenvalues, percent of the variance, and total variance accounted for by each canonical variable in Analyses 1, 2, and 3 for Sets 1 and 2.

Analysis	CV	Set 1			Set 2		
		E	% Var.	Cum. %	E	% Var	Cum %
1	1	5.365	76.0	76.0	3.943	70.4	70.4
	2	1.270	18.0	93.9	1.186	21.2	91.6
	3	.427	6.0	100.0	.466	8.3	99.9
	4	.001	.0	100.0	.005	.1	100.0
2	1	5.742	71.0	71.0	4.234	71.4	71.4
	2	1.296	16.0	87.0	1.044	17.6	89.0
	3	.878	10.9	97.8	.493	8.3	97.3
	4	.176	2.2	100.0	.159	2.7	100.0
3	1	9.162	78.9	78.9	8.247	73.7	73.7
	2	1.209	10.4	89.3	1.518	13.6	87.2
	3	.758	6.5	95.9	.916	8.2	95.4
	4	.479	4.1	100.0	.516	4.6	100.0

Note. CV = canonical variable; E = eigenvalue; % Var. = percent of total variance accounted for by each canonical variable; Cum. % = cumulative percentage of variation accounted for by the canonical variables.

Quantifying Speech Rhythm Abnormalities in the Dysarthrias

Julie M. Liss, Laurence White, Sven L. Mattys, Kaitlin Lansford, Andrew J. Lotto,
Stephanie M. Spitzer, and John N. Caviness
J Speech Lang Hear Res 2009;52;1334-1352; originally published online Aug 28,
2009;
DOI: 10.1044/1092-4388(2009/08-0208)

The references for this article include 5 HighWire-hosted articles which you can
access for free at: <http://jslhr.asha.org/cgi/content/full/52/5/1334#BIBL>

This information is current as of April 28, 2010

This article, along with updated information and services, is
located on the World Wide Web at:
<http://jslhr.asha.org/cgi/content/full/52/5/1334>

