Assessment of dysarthric speech through rhythm metrics

H. Dahmani [a,e,*], **S.-A. Selouani** [b], **D. O'shaughnessy** [a], **M. Chetouani** [c],
N. Doghmane [d]

[a] *INRS-EMT, Université du Québec, Canada*
[b] *Université de Moncton, Campus de Shippaga, Canada*
[c] *Université Pierre and Marie Curie, France*
[d] *Université Mokhtar Badji, Annaba, Algeria*
[e] *Université M'sila, M'sila, Algeria*

Abstract   This paper reports the results of acoustic investigation based on rhythmic classifications of speech from duration measurements carried out to distinguish dysarthric speech from healthy speech. The Nemours database of American dysarthric speakers is used throughout experiments conducted for this study. The speakers are eleven young adult males with dysarthria caused by cerebral palsy (CP) or head trauma (HT) and one non-dysarthric adult male. Eight different sentences for each speaker were segmented manually to vocalic and intervocalic segmentation (176 sentences). Seventy-four different sentences for each speaker were automatically segmented to voiced and non-voiced intervals (1628 sentences). A two-parameters classification related to rhythm metrics was used to determine the most relevant measures investigated through bi-dimensional representations. Results show the relevance of rhythm metrics to distinguish healthy speech from dysarthrias and to discriminate the levels of dysarthria severity. The majority of parameters was more than 54% successful in classifying speech into its appropriate group (90% for the dysarthric patient classification in the feature space (%V, ΔV)). The results were not significant for voiced and unvoiced intervals relatively to the vocalic and intervocalic intervals (the highest recognition rates were: 62.98 and 90.30% for dysarthric patient and healthy control classification respectively in the feature space (ΔDNV, %DV)).

© 2012 King Saud University. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

Dysarthria covers various speech impairments resulting from neurological problems and it probably represents a significant proportion of all acquired neurological communication disorders (as cited in Ziegler and von Cramon (1986)). These disorders are linked to the disturbance of brain and nerve stimuli of the muscles involved in the production of speech. Ultimately they induce disturbances in the strength, speed, range, tone, steadiness, timing, or accuracy of movements necessary for prosodically normal, efficient and intelligible speech (Liss et al., 2009; Yunusova et al., 2008). All types of dysarthria affect the articulation of consonants leading to slurring speech. Vowels may as well be distorted in very severe dysarthria. Rhythm troubles may be the most common characteristic of various types of dysarthria. Many studies state that most dysarthric patients

have slow speaking rates with long vowel and consonant segments as compared to standard control samples (Liss et al., 2009; Yunusova et al., 2008).

The present paper focuses on the assessment of rhythmic disturbance in dysarthria caused by cerebral palsy and head trauma. Cerebral palsy refers to a variety of developmental neuromuscular pathologies, occurring in three main forms: spastic, athetoid, and ataxic, associated with bilateral lesions of upper motor neuron pathways that innervate relevant cranial and spinal nerves. Dysarthria severity can be indexed in several ways, but quantitative measures usually consider features such as intelligibility and speaking rate. Disturbance of rhythm in the speech flow process is one of the important factors in dysarthric abnormalities (Liss et al., 2009). Even if the rhythm is identified as the main feature that characterizes dysarthria, assessment methods are mainly based on perceptual evaluation measures. Despite their numerous advantages that include the ease of use, low cost and clinicians' familiarity with related procedures, perceptual-based methods suffer a number of inadequacies and aspects that affect their reliability. These methods also lack evaluation protocols that may help standardization of judgments between clinicians and/or evaluation tools. Therefore, the aim of this work is to quantify rhythm abnormalities in the dysarthric speech by using the rhythm metrics developed recently in the language identification domain (Arvaniti and Rhythm, 2009).

This paper is organized as follows. Section 2 gives some definitions related to the rhythm metrics. Section 3 presents our method including the speech material, subjects and procedures used throughout experiments. In Section 4, we discuss the relevance of the rhythm metrics to assess the severity of dysarthrias. We describe the Gaussian Bayes classification system and its results in section 5. Section 6 concludes this paper.

2. Rhythm metrics

Rhythm metrics are based on acoustic measures of the duration of vocalic and consonantal intervals in continuous speech, they take into account variability in these durations, and they can be calculated in both raw and rate-normalized forms. A list of rhythm metrics used in our experiments is given at the end of this section. Grab and Low calculate durational variability in successive acoustic-phonetic intervals using Pairwise Variability Indices (PVI) (Grabe and Low, 2002). The raw Pairwise Variability Index (rPVI) is given in Eq. (1):

$$rPVI = \sum_{k=1}^{N-1} |d_k - d_{k+1}|/(N-1) \tag{1}$$

where dk is the length of the kth vocalic or intervocalic segment and N the number of segments.

A normalized version of the PVI index (noted nPVI) is defined by:

$$nPVI = \sum_{k=1}^{N-1} \left| \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right| /(N-1) \tag{2}$$

Ramus et al. (1999) based their quantitative approach of speech rhythm on purely phonetic characteristics of the speech signal. They measured vowel durations and the duration of intervals between vowels. They computed three acoustic correlates of rhythm from the measurements:

(a) %V: the proportion of time of vocalic intervals in the sentence;
(b) ΔV: the standard deviation of vocalic intervals;
(c) ΔC: the standard deviation of inter-vowel intervals.

Ramus et al. (1999) found that a combination of %V and ΔC provided the best acoustic correlate of rhythm classes. Our goal is to use these metrics in order to distinguish between the healthy and dysarthria speakers and to assess the speech intelligibility since the alterations of rhythm may also impact speech intelligibility. For each dysarthric sentence of each speaker, we have measured the durations of the vocalic, consonantal, voiced and unvoiced segments. In addition to the Vocalic-rPVI, Vocalic-nPVI, Intervocalic-rPVI, Intervocalic-nPVI, %V, ΔC, and ΔV, we computed the %DV, the duration of voiced intervals expressed in percent, ΔDV and ΔDNV, the standard deviation of voiced and non-voiced intervals, respectively.

3. Method

3.1. Speech material

Nemours is one of the few databases of recorded dysarthric speech. It contains records of American patients suffering different types of dysarthrias (Polikoff and Bunnell, 1999; James et al., 1996). The evaluation methodology followed in Nemours is inspired by the work of Kent et al. (1989). The test consists of a list of words from which four words are selected. The patient is supposed to listen to these words and repeat them aloud. The full set of stimuli consists of 74 monosyllabic names and 37 bi-syllabic verbs embedded in short nonsense sentences. Each Speaker pronounced 74 different sentences. Sentences have the following form: THE noun 1 IS verb-ING THE noun 2. The recording session was conducted by a speech pathologist considered as the healthy control (HC). The speech waveforms were sampled at 16 kHz and 16 bit sample resolution after low pass filtering at a nominal 7500 Hz cutoff frequency with a 90 dB/Octave filter.

3.2. Subjects

The speakers are eleven young adult males with dysarthria caused by cerebral palsy (CP) or head trauma (HT) and one non-dysarthric adult male (the experimenter). Seven speakers have CP, among whom three have CP with spastic quadriplegia and two have athetoid CP, and both have a mixture of spastic and athetoid CP with quadriplegia. The four remaining subjects are victims of head trauma. A two-letter code was assigned to each patient: BB, BK, BV, FB, JF, KS, LL, MH, RK, RL and SC. Thanks to the Frenchay dysarthria assessment scores (see Table 1 and (James et al., 1996; Enderby and Pamela, 1983)), the patients can be divided into three subgroups: one mild, including subjects FB, BB, MH and LL; the second subgroup includes the subjects RK, RL, and JF and the third is severe and includes subjects KS, SC, BV, and BK. The perceptual data and the speech assessment did not take into consideration the too severe case (patient KS) and the too mild case (patient FB).

4. Experiments and results

The mean and the standard deviation of vocalic and consonantal interval durations are given in the Fig. 1. These measures

Table 1 Frenchay dysarthria assessment scores of dysarthric speakers of Nemours database (James et al., 1996).

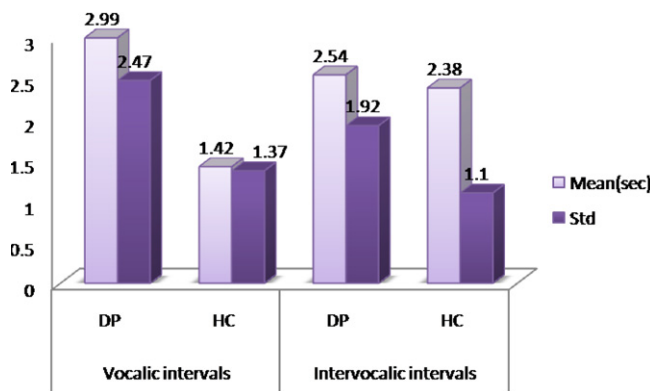| Patients	Level of severity (%)	Intelligibility
KS	–	1
SC	49.2	3
BV	42.5	3
BK	41.8	3
RK	32.4	2
RL	26.7	4
JF	21.5	4
LL	15.6	6
BB	10.3	8
MH	7.9	6
FB	7.1	–

Figure 1 Mean and standard deviation of consonantal and vocalic intervals durations.

confirm clearly that the durations of both intervals are greater for Dysarthric Patients (DP) than the Healthy Control (HC). This result leads us to carry out a set of experiments that aim at creating bi-dimensional maps (One metric with respect to another) that might be useful for analyzing the relevance of these metrics to both clinical practice and research purposes. Besides this, the one-way analysis of variance (ANOVA) is also performed in order to assess the statistical significance of rhythm metrics for dysarthric speech classification.

4.1. Investigation of %V, ΔV and ΔC

The one-way ANOVA is carried out to determine if the metrics demonstrated significant group differences. The main effect of group differences was statistically significant for ΔC, ΔV and %V, ($F_{(1,174)} = 67.19$, $p < 0.000$, $F_{(1,174)} = 65.84$ $p < 0.000$ and $F_{(1,174)} = 6.125$, $p = 0.012$, respectively). Although, we note that %v is less significant than ΔV and ΔC.

Fig. 2 shows the distribution of DP and HC along the %V (x- axis) and ΔV (y-axis) dimensions. This (%V, ΔV) feature space shows that except for BV the distribution matches very well the Frenchay dysarthria assessment scores. The DP patients: KS, BK, SC, that belongs to the most severe cases are characterized by a relatively low %V. According to this distribution RL can be considered a severe case. The realizations of the healthy control are well regrouped around the nominal values of these parameters but we can note that only MH coincides with HC's group.
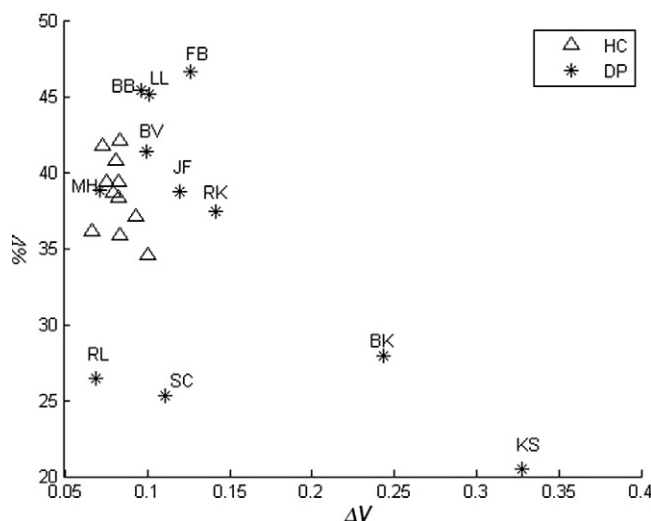

Figure 2 Distribution of DP (Dysarthric Patients) and HC (Healthy Controls) into the two-dimensional (%V, ΔV) feature space.
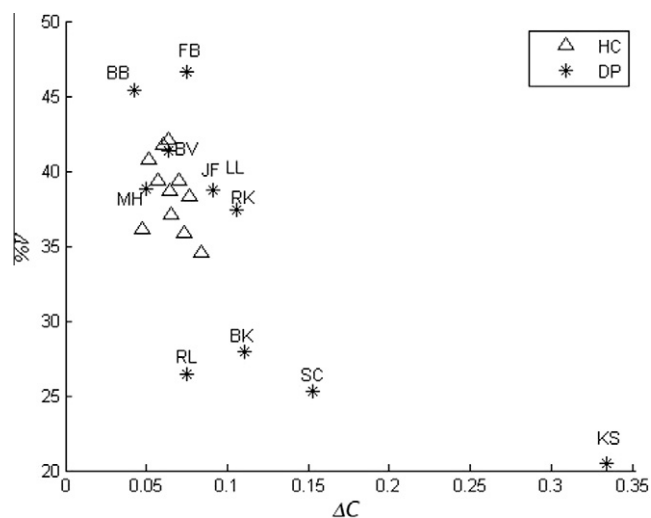

Figure 3 Distribution of DP (Dysarthric Patients) and HC (Healthy Controls) into the two-dimensional (ΔC, %V) plan.

Fig. 3 gives the 2-D distribution of DP and HC into the (%V, ΔC) feature space. The DPs, particularly the most severe cases, can also be easily distinguished from the HC despite that BV who is a severe patient is among the HC.

4.2. Investigation of n-PVI and r-PVI Metrics

The ANOVA analysis reveals that the main effect of group for Vocalic-rPVI and Vocalic-nPVI was statistically significant with $F_{(1,174)} = 60.59$, $p < 0.001$, and $F_{(1,174)} = 1.006$, $p < 0.001$, respectively. The main effect of group for intervocalic-rPVI and intervocalic-nPVI metrics was also-statistically significant with $F_{(1,174)} = 20.156$, $p < 0.001$, and $F_{(1,174)} = 59.231$, $p < 0.001$, respectively.

The relevance of the Pairwise Variability Index to assess dysarthrias was investigated through bi-dimensional representations. In Fig. 4, the (vocalic-nPVI, intervocalic nPVI) feature
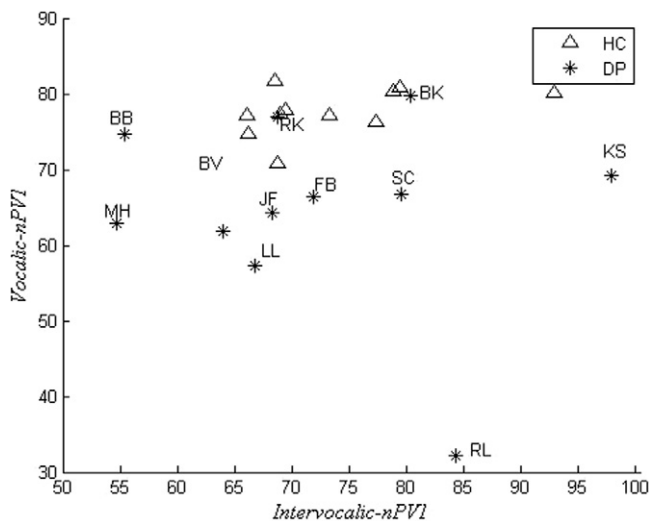
Figure 4 Dysarthric and healthy subjects represented into the (intervovalic nPVI, vocalic nPVI) feature space.
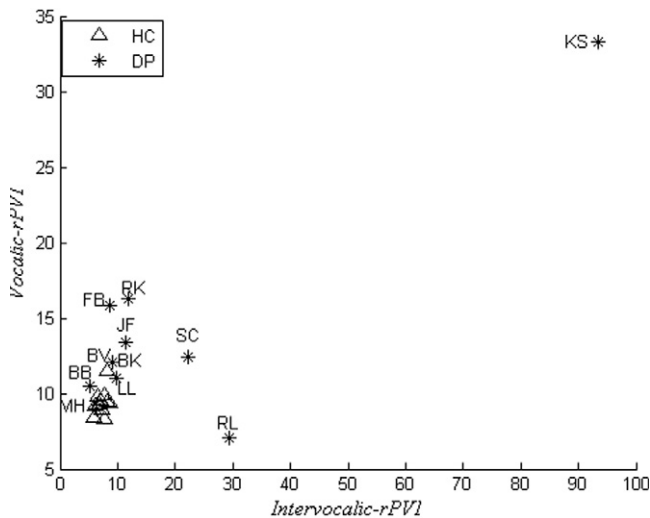

Figure 5 Dysarthric and healthy subjects represented into the (intervovalic rPVI, vocalic, rPVI) feature space.

space shows that HC are relatively well grouped around higher vocalic-nPVI values and mid intervocalic-nPVI, while DPs are more scattered in this space. The particular position of RL is quite surprising since the Frenchay test categorizes it in the mild category.

Fig. 5 shows the 2-D distribution of DP and HC into the (vocalic-rPVI, intervocalic-rPVI) plan. We can notice that BV who is considered a severe case is always close to HC and mild DP. FB the mildest, was relatively far to HC. Therefore, this representation seems inadequate for a good discrimination of the subjects. Suitability of using the most severe case, KS, is well discriminated by the two representations.

4.3. Investigation of voiced and voiceless dysarthric speech

The Anova test reveals that the main effect for %DV, ΔDNV was significant ($F(1, 1615) = 63.597$, $p < 0.001$, $F(1, 1615) = 63.597$, $p < 0.001$, respectively).
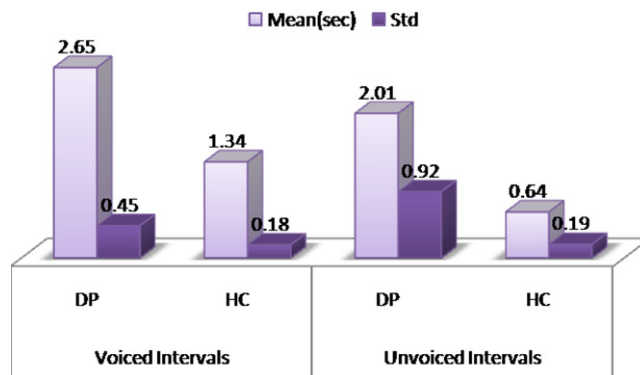

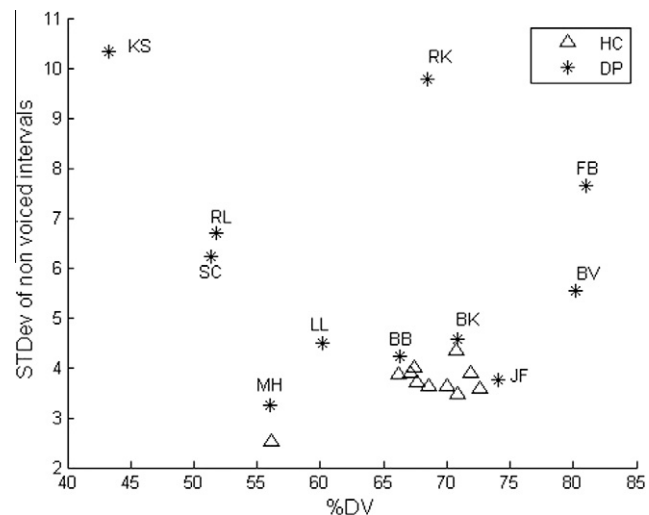Figure 6 The mean and standard deviation computed for voiced and voiceless intervals duration.

Figure 7 Dysarthric and healthy subjects represented in the (%DV Δdnv) feature space.

It was shown in Ackermann and Hertrich (1994) and in Platt (1980) that cerebral palsied individuals lack articulatory precision. Athetoid cerebral palsied patients are likely to make imprecise articulation of word initial consonants. In addition, it was shown that involuntary breathing and jaw movement affect both consonant and vowel production due to distortion in the place and manner of articulation and in formant patterns (Bon and Horowitz, 1993). In order to determine to what extent the decreased intelligibility was caused by difficulties with voiced-voiceless distinctions (laryngeal timing), the voiced and voiceless intervals of dysarthric speech were analyzed. The total of 74 sentences uttered by each dysarthric speaker and those of the healthy control were segmented and labeled to voiced and voiceless intervals. The duration measures carried out on both voiceless and voiced intervals reveal noticeable differences between HC and DPs. As shown in Fig. 6, dysarthric speakers tend to produce lengthened voiceless segments with higher values of standard deviation. The duration of sentences repeated by KS (the most severe case) was far superior to other DP durations.

In Fig. 7, we have plotted a 2-D distribution of DP and HC into the plan formed by the standard deviation of non-voiced

intervals and the percentage of voiced intervals duration ($\Delta DNV, \%DV$). We can observe from Fig. 5 a random distribution of the DP while the HC are well regrouped. The most severe cases are positioned far from HC with relatively higher ΔDNV. The rest of DPs are close to HC. In all the 2-D representations we performed, we noted that BV, whose Frenchay and intelligibility scores are 57.5 and 3, respectively, is always positioned close to the HC. Indeed, on examining the speech of BV and FB patients, we noted that the speed of BV speech was quite normal and almost intelligible but with nasality. We have also noted that FB is the mildest case but he is not the closest DP to HC. In fact his speech is very intelligible but his speech rate is very slow.

5. Classification system

A Gaussian Bayes classification was used to evaluate the features with respect to their discriminatory power. It is a simple method for supervised classification based on the use of Bayes' theorem. In a Gaussian model, a set of data is characterized by the mean and covariance for each class within the data along a number of dimensions. For each new data point, we calculate the probability that that point came from each class; the data point is then assigned to the class which gave the highest probability.

A classification in two dimensions was carried out. In our case, we have considered the combinations of two parameters given by Tables 2 and 3. A closed test that involves training and testing on the same data was utilized.

Table 2 shows the results for the vocalic and intervocalic segmentation. We can see that the feature space of ($\%V, \Delta V$) gives the best separation score with overall rate of correct classification 73.85% (53.4% for DP and 94.3 for HC) (we can see

that Fig. 2 gives the best separation between the two groups DP and HC). But also the feature space of the normalized Pairwise Variability Index (vocalic-nPVI, intervocalic-nPVI) whose 72.15% overall correct separation of the dysarthric patients from healthy control (65.9% for DP and 78.4 for HC) is a very encouraging score in spite of the closed test and the limited size of data.

Table 3 illustrates the results for the voiced and unvoiced segmentation. All in all, we note that the results are less important than given by Table 2. The most important result is for the bidimensional distribution ($\%DV, \Delta DNV$) (62.98% for DP and 90.3% for HC) (we can verify this result by Fig. 7). The worst result is for the normalized Pairwise Variability Index (nPVI)).

To examine the weakness of the parameters using to separate the DP and HC groups, we have performed the classification in two dimensions to the severity levels. The severe patients (KS, SC, BV and BK) designed by SP, the moderate patients (JF, RL and RK) by MOP and mild patients (FB, MH, BB and LL) by the two letters MP. The results for both vocalic and voiced segmentation are given by the Tables 4 and 5.

As expecting especially for the feature space ($\%V, \Delta V$), the problem was in the separation of the mild patients from HC because of its height rate of misclassification 68.8% as belonging to HC and 3.1% and 0% as belonging to MOP and as SP groups respectively. With the normalized pairwise parameters, the severe patients were almost misclassified as HC and MP patients (62.5% as HC and 31.3% as MP).

For the voiced and unvoiced segmentation, the common result is the height rate of misclassification of the MP patient as HC for all two feature space classification. We note 68.8% for ($\%DV, \Delta DNV$), 80.3 for ($\%VO, \Delta VO$), 82.7 for (Voiced-rPVI, unvoiced-rPVI) and the highest rate: 89.2 for (Voiced-nPVI, unvoiced-nPVI).

Table 2 The summary of classification results for the vocalic and intervocalic intervals.

	(Vocalic-nPVI, intervocalic-nPVI)		(Vocalic-rPVI, intervocalic-rPVI)		(% V, ΔV)		(% V, ΔC)	
	DP	HC	DP	HC	DP	HC	DP	HC
DP	65.9	34.1	42.05	57.95	53.4	46.6	48.86	51.14
HC	21.1	78.4	2.3	97.7	5.7	94.3	3.4	96.6

Table 3 The summary of classification results for the voiced and unvoiced intervals.

	(Voiced-nPVI, unvoiced-nPVI)		(Voiced-rPVI, unvoiced-rPVI)		(%VO, ΔVO)		(%DV, ΔDNV)	
	DP	HC	DP	HC	DP	HC	DP	HC
DP	48.3	51.7	46.6	53.4	46.3	53.7	64.3	35.7
HC	26.5	73.5	8.5	91.5	6.6	93.4	9.8	90.2

Table 4 The summary of classification results for severity levels with the vocalic and intervocalic intervals.

	(Vocalic-nPVI, intervocalic-nPVI)				(Vocalic-rPVI, intervocalic-rPVI)				(%V, ΔV)				(%V, ΔC)			
	MP	MOP	SP	HC	MP	MOP	SP	HC	MP	MOP	SP	HC	MP	MOP	SP	HC
%	37.5	25.0	3.1	55.7	9.3	33.3	65.6	94.3	28.1	58.3	56.3	89.8	21.9	41.7	62.5	94.32

Table 5 The summary of classification results for severity levels with the voiced and unvoiced intervals.

	(Voiced-nPVI, unvoiced-nPVI)				(Voiced-rPVI, unvoiced-rPVI)				(%VO, ΔVO)				(%DV, ΔDNV)			
	MP	MOP	SP	HC	MP	MOP	SP	HC	MP	MOP	SP	HC	MP	MOP	SP	HC
%	0.3	0.5	11.6	95.1	6.4	18.55	43.5	93.9	12.54	19.0	48.0	94.0	20.7	29.0	52.0	93.0

6. Conclusion

In this paper, we presented a rhythm-based method for the assessment of dysarthric speech. Rhythm metrics based on durational characteristics of vocalic and intervocalic intervals and Pairwise Variability Index using with both their raw and normalized measures are used for this purpose. We found that these metrics are not very promising to express the severity level of the dysarthria impairment, while noting that mild cases are difficult to classify despite the general trend of being classified as healthy cases.

For combinations of parameters chosen in this article, we can infer that the pairs (%V, ΔV) and (Vocalic-rPVI, intervocali-i-rPVI) were the best. The results for the voiced and unvoiced intervals were not relatively important but the ease of segmentation and extraction of parameters remain as major advantages. Therefore, we believe that the results for both types of segmentation can be improved by using more rhythm features and thus using classifications in higher dimensional spaces.

This study examined variation in rhythm metrics by focusing on the differences between healthy and dysarthric speakers. The experiment looked at the realization of the duration contrast among these speakers. The main result is that rhythm metrics are sensitive to differences between groups of dysarthric speakers. The methodological implication of this result is that for some cases the subjective Frenchay test may incorrectly categorize some subjects. We suggest adding the exploit of rhythm metrics in the design of dysarthria assessment.

The future work should focus on the need to automate the measurements and to use more metric features with more speech samples to increase the reliability of the results. We also need to focus on studying more effective rhythm features to improve the severity levels and speaker recognition rate in the further.

Appendix A. Speech material for set 1: dysarthric talkers are identified by a two-letter code (e.g., bb.wav, rk.wav, sc.wav, etc.) And the parallel productions from a normal adult male talker have the code JP prepended to the filename (e.g., jpbb1.wav, jpbb2.wav, etc.)

BB1 (JPBB1): The bash is pairing the bath
BB2 (JPBB2): The bad is sleeping the bin
BB3 (JPBB3): The two is weeping the bit
BB4 (JPBB4): The kong is licking the chin
BB5 (JPBB5): The bat is shooing the chew
BB6 (JPBB6): The pat is wearing the sue
BB7 (JPBB7): The thin is surging the vat
BB8 (JPBB8): The pin is knowing the cop

RL1 (JPRL1): The con is bearing the butt
RL2 (JPRL2): The bait is waking the bet
RL3 (JPRL3): The rot is weeping the boat
RL4 (JPRL4): The back is heaping the yacht
RL5 (JPRL5): The com is licking the vat
RL6 (JPRL6): The lot is surfing the fate
RL7 (JPRL7): The phase is chewing the dive
RL8 (JPRL8): The faith is sitting the fade

BK1 (JPBK1): The sin is sitting the who
BK2 (JPBK2): The goo is surfing the batch
BK3 (JPBK3): The Bert is sinning the die
BK4 (JPBK4): The fife is waking the bad
BK5 (JPBK5): The watt is waning the dive
BK6 (JPBK6): The back is stewing the com
BK7 (JPBK7): The bin is pairing the tin
BK8 (JPBK8): The coo is singing the thin

BV1 (JPBV1): The fat is surging the bat
BV2 (JPBV2): The chew is waking the fine
BV3 (JPBV3): The fife is owing the cop
BV4 (JPBV4): The bit is sipping the mat
BV5 (JPBV5): The inn is shooing the din
BV6 (JPBV6): The two is knowing the moo
BV7 (JPBV7): The faith is mowing the Jew
BV8 (JPBV8): The thin is sinning the knew

FB1 (JPFB1): The yacht is lifting the bin
FB2 (JPFB2): The dial is suing the die
FB3 (JPFB3): The five is weeping the butt
FB4 (JPFB4): The bathe is going the thin
FB5 (JPFB5): The shoe is pairing the beet
FB6 (JPFB6): The boat is heaping the fat
FB7 (JPFB7): The dive is reaping the bat
FB8 (JPFB8): The bite is listing the bet

JF1 (JPJF1): The fay is sitting the Bert
JF2(JPJF2): The thin is shooing the dew
JF3 (JPJF3): The tin is bearing the bit
JF4 (JPJF4): The yacht is knowing the cop
JF5 (JPJF5): The badge is weighing the bat
JF6 (JPJF6): The coo is stewing the fake
JF7 (JPJF7): The com is living the base
JF8 (JPJF8): The beet is mowing the butt

KS1 (JPKS1): The fife is lifting the boat
KS2 (JPKS2): The bake is shooing the five
KS3 (JPKS3): The bathe is reaping the dial
KS4 (JPKS4): The watt is tearing the phase
KS5 (JPKS5): The con is licking the bash
KS6 (JPKS6): The shoe is suing the goo
KS7 (JPKS7): The back is stewing the thin
KS8 (JPKS8): The vat is living the Jew
LL1 (JPLL1): The fate is stewing the sue
LL2 (JPLL2): The gin is going the dew
LL3 (JPLL3): The con is knowing the cob

LL4 (JPLL4): The back is sitting the tin
LL5 (JPLL5): The lot is surging the inn
LL6 (JPLL6): The fife is licking the bin
LL7 (JPLL7): The chin is daring the bathe
LL8 (JPLL8): The rot is sweeping the dial
MH1 (JPMH1): The cop is suing the badge
MH2 (JPMH2): The fate is serving the phase
MH3 (JPMH3): The bathe is shooing the rot
MH4 (JPMH4): The goo is wearing the bet
MH5 (JPMH5): The beet is lifting the lot
MH6 (JPMH6): The vat is weeping the five
MH7 (JPMH7): The bite is knowing the batch
MH8 (JPMH8): The boat is owing the boot

RK1 (JPRK1): The dew is bearing the fight
RK2 (JPRK2): The bet is going the thin
RK3 (JPRK3): The shin is shooing the yacht
RK4 (JPRK4): The rot is wading the coo
RK5 (JPRK5): The fin is weighing the bait
RK6 (JPRK6): The dial is singing the zoo
RK7 (JPRK7): The chin is surging the fay
RK8 (JPRK8): The com is pairing the kong

SC1 (JPSC1): The bin is sleeping the con
SC2 (JPSC2): The fat is reaping the dime
SC3 (JPSC3): The badge is waking the bad
SC4 (JPSC4): The fin is shooing the face
SC5 (JPSC5): The sue is sipping the sin
SC6 (JPSC6): The batch is listing the lot
SC7 (JPSC7): The yacht is living the gin
SC8 (JPSC8): The zoo is daring the bash

References

Ackermann, H., Hertrich, I., 1994. Speech rate and rhythm in cerebellar dysarthria: an acoustic analysis of syllable timing. Folia Phoniatrica 46, 70–78.

Arvaniti, A., Rhythm, A., 2009. Rhythm, timing and the timing of rhythm. Phonetica 66, 46–63.

Bon, S.K., Horowitz, D.M., 1993. A statistical causal model for the assessment of dysarthric speech and the utility of computer-based speech recognition. IEEE Transactions on Biomedical Engineering 40 (12).

Enderby, P., Pamela, M., 1983. Frenchay Dysarthria Assessment. College Hill Press.

Grabe, E., Low, E.L., 2002. Durational variability in speech and the rhythm class hypothesis. In: Papers in Laboratory Phonology 7.

James, B., Peters, S.M., Leonzio, J.E., Bunnell, H.T., 1996. The Nemours database of dysarthric speech. In: Proceedings of the Fourth International Conference on Spoken Language Processing, October 3–6, Philadelphia, PA, USA.

Kent, R.D., Weismer, G., Kent, J.F., Rosenbek, J.C., 1989. Toward phonetic intelligibility testing in dysarthria. Journal of Speech and Hearing Disorders 54, 482–499.

Liss, J.M., White, L., Mattys, S.L., Lansford, K., Lotto, A.J., Spitzer, S., Caviness, J.N., 2009. Quantifying speech rhythm abnormalities in the dysarthrias. Journal of Speech Language and Hearing Research 52, 1334–1352.

Platt, L.J., 1980. Dysarthria of adult palsy. Journal of Speech Hearing Research, 28–55.

Polikoff, J.B., Bunnell, H.T., 1999. The Nemours database of dysarthric speech: a perceptual analysis. In: The XIVth International Congress of Phonetic Sciences (ICPhS), San Francisco, USA.

Ramus, F., Nespor, M., Mehler, J., 1999. Correlates of linguistic rhythm in the speech signal. Cognition 73, 265–292.

Yunusova, Y., Weismer, G., Westbury, J.R., Lindstrom, M., 2008. Articulatory movements during vowels in speakers with dysarthria and normal controls. Journal of Speech, Language, and Hearing Research 51 (3), 596–611.

Ziegler, W., von Cramon, D., 1986. Disturbed coarticulation in apraxia of speech: acoustic evidence. Brain and Language 29, 34–47.