ORIGINAL RESEARCH

DYNAMICAL CHARACTERISTICS OF SPEECH APPARATUS IN HUNTINGTON'S DISEASE

Jan Hlavnička

Dept. of Circuit theory, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic

Abstract

This paper presents new approach to measuring dynamical characteristics of the articulation and phonation apparatus in Huntington's disease (HD). This study utilizes the rhythm task in unusual way. Each repetition of the syllable is explored as unique measurement of dynamical characteristics of speech apparatus. The evaluation of dynamical characteristics was implemented into fully automatic algorithm. The database is gender balanced and consists of 43 Czech native participants including 20 healthy control (HC) and 23 HD subjects. The evaluation is based on recently designed segmentation method with very high accuracy of syllable detection and set of 6 descriptive features. All features of the set showed significant statistical difference ($p<0.05$) between HC and HD group. Longer duration of syllables and longer volume increase time reflect the decline of coordination of vocal and articulatory muscles. Higher instability in volume increase time and instability of reproduction are related to ataxic movements of vocal muscles in HD. Instability of duration of syllables and pauses corresponds with all motor manifestations in HD include vocal, articulatory and respiratory system dysfunction.

Keywords

Huntington's disease, syllable detection, rhythm test, speech disorders, speech assessment.

Introduction

Huntington's disease (HD) is an autosomal dominantly inherited neurodegenerative disorder caused by frequent repetition of an unstable $(CAG)_n$ trinucleotide repeats on chromosome 4p16.3 [1]. The onset age variates from 20 to 40 years, but HD typically manifests in third decade of life. The neurodegenerative process leads to progression of uncontrolled movements, emotional problems, and loss of cognition ability over 10 to 20 years until death [2]. A person with HD has a 50% chance to pass his huntingtin gene (HTT) to his descendants. A genetic test can confirm HTT, but HTT gives just the information about predispositions nor the onset age.

Human speech is a very complex movement and even subtle motor, cognition or control dysfunction causes characteristic dysarthria. Assessment of dysarthria could be a very sensitive and efficient method for screening potential neurodegenerative disorders e.g. Parkinson's disease [3], Alzheimer's disease [4], multiple sclerosis [5], and of course HD, and many others. Speech in HD is affected into hyperkinetic dysarthria with influenced phonation, timing, prosody, and with occurrence of articulation breakdowns [6].

Steadiness of function of vocal muscles in sustained vowel phonation provides significant features suitable even for classification of preclinical HD [7]. Timing, prosody and articulation of speech were also previously researched. This study aims with the new measurement of dynamic function of vocal muscles and articulation apparatus in HD.

The rhythm task consist of perpetuate articulation of syllable /Pa/. It is common to evaluate performance of the rhythm task in terms of rhythm stability and acceleration to evaluate timing of speech in HD [8]. This study utilizes the rhythm task in very unusual way. Each repetition of the syllable is explored as unique measurement of dynamical characteristics of the articulation and phonation apparatus. The present approach to the rhythm task was never previously published. The proposed assessment of dynamical

characteristics was designed into fully automatic method.

Methods

Subjects

The database consists of 43 Czech native participants originally recruited for previous study [9]. The rhythm task was previously examined in terms of rhythm stability and acceleration [8]. The group of 20 HD patients (9 men, 11 women) was clinically and genetically verified with mean disease duration 6.4± SD 3.0 (2-13) years and disease severity 23.6 ± 11.8 (3-42) of Unified Huntington's Disease Rating Scale (UHDRS) [10]. Motor subscore ranges from 0 to 124, where higher score means more severe disability. The group of 23 healthy control (HC) subjects (12 men, 11 women) was at mean age 64.4 ± 10.5 (41-81) years.

HD patients were treated by benzodiazepines, antipsychotics, amantadine and antidepressants, in monotherapy or in various combinations. The input condition for HD was the ability to sustain prolonged phonation for at least 6 seconds and to perform at least 20 syllables in sequence. Each participant provided written, informed consent. The study was approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic.

Recordings

Speech data were recorded in a quiet room with a low ambient noise level using a head-mounted condenser microphone (Bayerdynamic Opus 55, Heilbronn, Germany) situated approximately 5 cm from the mouth. Recordings were sampled at 48 kHz with 16-bit resolution.

All participants were asked to repeat the syllable /Pa/ at least 20 times at a comfortable, self-determined, steady pace without acceleration or deceleration. Each subject performed the task two times.

Segmentation

Each recording was segmented using recently designed method with very high accuracy 99.6±2% of syllable detection [8]. The method includes sensitive syllable detection based on adaptive recognition and algorithm of false positives rejection based on cluster analysis. The method is shown in the flow diagram 1.A.

Classification

We decimated the signal into a sampling frequency of 10 kHz, because frequencies above 5 kHz are redundant in the precise detection of syllables. Signal was described using 12 Mel-frequency cepstral coefficients (MFCC) inside a sliding window of 10 ms length, 3 ms step and hamming weighting. Most information about the signal origin (syllable / pause) is contained in the first three MFCC. High sensitivity is required, therefore Syllables were classified using k-mean algorithm inside a recognition window of 4 s length and 800 ms step. Classification is illustrated in figure 1.B.

Components were identified with syllables in condition of higher mean of the first MFCC (related to power). The decision smoothing consisted of median filter of the 5th order and shorter pulses (< 30 ms) rejection and shorter pauses (< 80 ms) rejection.

Syllable parameterization

Each syllable was parametrized into vector of means of first three MFCC. All parametrized syllables form the parametric space F.

Outlier detection

Outlier detection presupposes normal distribution of syllable parameters in the space F. The number of observations is relatively low, therefore unorthodox algorithm for outlier detection was created.

Means and variances of normal distributions of syllables and false detected syllables (mostly loud respirations) are unique for each speaker. Let X be the

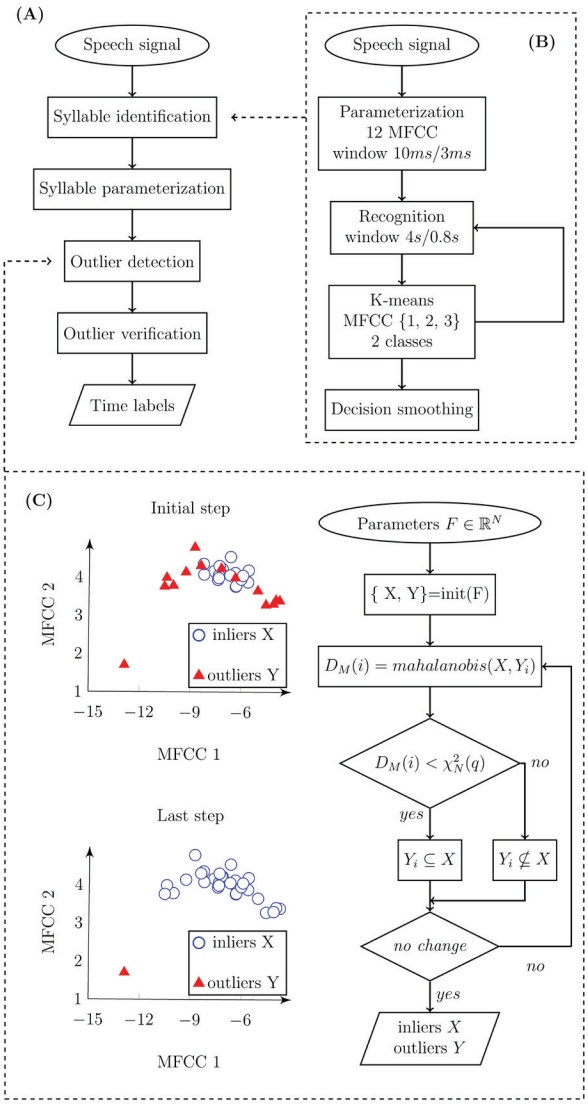

Fig. 1: The flow diagram of the automatic syllable segmentation method [8].

set of inliers and Y be the set of outliers. The distance between outlier Y_i to the set of inliers X with the respect to the variance and the mean of inliers X was measured using Mahalanobis distance:

$$DM(i) = \sqrt{(Y_i - \mu)^T S^{-1} (Y_i - \mu)}, \quad (1)$$

where $DM(i)$ is Mahalanobis distance from the outlier Y_i to the inliers distribution X, Y_i is the observation of the outlier, μ is mean of the inliers distribution X, and S is the covariance matrix of the inliers distribution X. Mutual Mahalanobis distances of any N-dimensional normal distribution form χ^2 distribution with N-degrees of freedom. It is routine to look for outliers in some high quantile $\chi^2_N(q)$, typically $q = 0.975$. However, outlier identification in a single step at such a small number of observations does not contribute sufficient sensitivity. Therefore, we designed identification procedure, which involves the gradual recognition of outliers.

The initial set of inliers X included observations Y_i with condition:

$$DM(Y_i) < \chi^2_N(q), \quad (2)$$

where DM was measured and thresholded within all observations in quantile $q = 0.3$. We determined the most probable members of X. In recognition steps, we measured the $DM(Y_i)$ between the each member of Y and the set of X. False positives occur mostly above the estimated quantile $q = 0.5$. Observations in accordance with this condition belongs to set of X. The recognition step was repeated until no change in set X was evidenced. The process of outlier identification is shown in figure 1.C.

Outlier verification

Occasionally, some loud or quiet syllables have different spectral envelope and can act as outliers. But its high energy differs them from other outliers. The verification is necessary. The speech signal was filtered using a Chebyshev filter of the 5th order in 100-500 Hz band pass. We parametrized the filtered signal into power inside a sliding window of 10 ms length, 3 ms overlap and hamming weighting. Each syllable P_X was described as the mean of power and each outlier P_Y to the maximum value of the power. Subsequently, the outlier Y_i was rejected on 95% population level of one-sided Chebyshev's inequality:

$$P(Y_i) > E(P_X) - 4 \cdot \sigma(P_X), \quad (3)$$

where E denotes mean and σ is standard deviation.

Time labels

Syllables were described into labels as time of each syllable onset, time of highest filtered energy peak of each syllable and time of occlusion of each syllable.

Features

Rhythm features are standard result of the rhythm test. This set of features focuses on non-rhythmical features to describe dynamic function of vocal muscles and articulation apparatus in HD. Features are based on time labels of automatic segmentation.

- Duration of syllables (DS)

Mean length of syllables.

- Stability of duration of syllables (SS)

Standard deviation of length of syllables.

- Stability of duration of pauses (SP)

Standard deviation of length of pauses.

- Volume increase time (VIT)

Mean time interval between syllable onset and time of syllable loudness maxima.

- Stability of volume increase time (SIT)

Standard deviation of time interval between syllable onset and time of syllable loudness maxima.

- Stability of reproduction (SR)

Each syllable was parametrized into power inside a sliding window of 10 ms length, 3 ms overlap and hamming weighting. The power envelope in logarithmical scale was transformed using discrete cosine transformation. SR was computed as sum of standard deviations of 2nd to 6th discrete cosine transformation coefficients.

Each speaker obtained final feature assessment calculated as mean value of each feature within two rhythm tasks.

Statistics

Feature significance was evaluated by the unpaired two-tailed t-test. Normality of distribution of features was tested using one-sample Kolmogorov–Smirnov test.

Results

Each feature of the set showed significant difference ($p<0.05$) between HC and HD groups, confirming the downgrade of vocal and articulatory function across all measured dimensions. DS showed significant difference ($p<0.01$). Features SS, SIT, SP and SR differentiated groups with significance ($p<0.001$). Results of all features is summarized in Tab. 1.

Discussion

The current study presents fully automatic method for measurement of dynamical characteristics of phonation and articulation apparatus in the rhythm test. Rhythm task is usually used for measurement of speed and acceleration of rhythm. Proposed method investigate the dynamical function of speech apparatus during repeated syllables. Longer duration of syllables shows deceleration of vocal muscles and longer volume increase time evidence decline in coordination of articulatory muscles. Higher instability in volume increase time and instability of reproduction are related to ataxic movements of vocal muscles in HD. Instability of duration of syllables and pauses are in consequences of all motor manifestations of HD, include vocal, articulatory and respiratory system dysfunction.

Tab. 1: Feature characteristics and statistical significance within the groups HC and HD.

	HC		HD		
	Mean	SD	Mean	SD	t-test
DS	0.16	0.04	0.20	0.06	p=0.003
SS	0.02	0.01	0.05	0.03	p<0.001
VIT	0.06	0.01	0.07	0.03	p=0.013
SIT	0.02	0.01	0.04	0.01	p<0.001
SP	0.04	0.02	0.14	0.08	p<0.001
SR	13.82	3.8	20	5.88	p<0.001

Acknowledgment

This research was supported by Czech Science Foundation (GACR 102/12/2230) and Grant Agency of the Czech Technical University in Prague (SGS15/199/OHK3/3T/13).

References

[1] Kremer, B., Goldberg, P., Andrew S.E., Theilmann, J., Telenius, H., et al. *A worldwide study of the Huntington's disease mutation: The sensitivity and specificity of measuring CAG repeats*. The New England Journal of Medicine, 1994, vol. 330, p. 1401–1406.

[2] Martin, J.B., and Gusella, J.F. *Huntington's disease. Pathogenesis and management.* The New England Journal of Medicine, 1986, vol. 19, p. 1267-1276.

[3] Harel, B.T., Cannizzaro, M.S., Cohen, H., Reilly, N., Snyder, P.J. *Acoustic characteristics of Parkinsonian speech: A potential biomarker of early disease progression and treatment.* Journal of Neurolinguistics, 2004, vol. 14, p. 439-453.

[4] Ballard, K.J., Savage, S., Leyton, C.E., Vogel, A.P., Hornberger, M., et al. *Logopenic and Nonfluent Variants of Primary Progressive Aphasia Are Differentiated by Acoustic Measures of Speech Production*. PLOS ONE, 2014, vol. 9, e89864.

[5] Tjaden, K., and Willding, G. *Speech and pause characteristics associated with voluntary rate reduction in Parkinson's disease and Multiple Sclerosis*. Journal of Communication Disorders, 2011, vol. 44, p. 655-665.

[6] Duffy, J.R. *Motor Speech Disorders: Substrates, Differential Diagnosis and Management, 2nd edition*. Mosby, 2005, p. 592.

[7] Rusz, J., Saft, C., Schliegel, U., Hoffman R., Skodda, S. *Phonatory Dysfunction as a Preclinical Symptom of Huntington Disease*. PLOS ONE, 2014, vol. 9, p. 1-7.

[8] Rusz, J., Hlavnička, J., Čmejla, R., and Růžička, E. *Automatic evaluation of speech rhythm instability and acceleration in dysarthrias associated with basal ganglia dysfunction.* Frontiers in Bioengineering and Biotechnology, 2015, vol. 3, p. 1-11.

[9] Rusz, J., Klempíř, J., Tykalová, T., Baborová, E., Čmejla, R., Růžička, E., and Roth, J. *Characteristics and occurrence of*

speech impairment in Huntington's disease: possible influence of antipsychotic medication. Journal of Neural Transmission, 2014, vol. 121, p. 655-664.

[10] Huntington Study Group. *Unified Huntington's Disease Rating Scale: reliability and consistency.* Movement disorders, 1996, vol. 11, p. 136-142.

Ing. Jan Hlavnička
Department of Circuit Theory
Faculty of Electrical engineering
Czech Technical University in Prague
Technická 2, 166 27 Praha

E-mail: hlavnjan@fel.cvut.cz