International Journal of the Physical Sciences Vol. 7(3), pp. 394 - 399, 16 January, 2012
Available online at http://www.academicjournals.org/IJPS
DOI: 10.5897/IJPS11.1561
ISSN 1992 - 1950 © 2012 Academic Journals

Full Length Research Paper

The acute hypoxia indication by the dissonant intervals of the speech signal

Zoran N. Milivojević[1], Marina Milivojević[2], Darko Brodić[3]* and Dragan R. Milivojević[4]

[1]Techical College Niš, Aleksandra Medvedeva 20, 18000 Niš, Serbia.
[2]University of Niš, Medical Faculty, Bulevar dr Zorana Đinđića 81, 18000 Niš, Serbia.
[3]University of Belgrade, Techical Faculty in Bor, Bor, V.J. 12, 19210 Bor, Serbia.
[4]Institute for Mining and Metallurgy, Zeleni bulevar bb, 19210 Bor, Serbia.

Accepted 14 December, 2011

The paper analyzes the effect of acute hypoxia to dissonant intervals at the speech signal. It presents the factors that lead to the hypoxia as a consequence of the oxygen concentration decrease in the tissue. Furthermore, a definition of dissonant spectrum ranges of the speech signal is determined. It is made in relation to the fundamental frequency. It stands as tones' F#, B and C# in relation to C. For the experiment, the articulation of vowels from the group of persons was recorded. It is performed at 200, 800, 1400 and 1900 m altitudes above the sea level. The result of this speech signal testing is summarized and shown in graphs. Furthermore, the dissonant coefficient $k_{F\#}$, k_B and $k_{C\#}$, have also been determined. Their values indicate the degree of the acute hypoxia.

Key words: Speech signal, fundamental frequency, dissonant frequency, acute hypoxia.

INTRODUCTION

Hypoxia represents a condition of decreased oxygen concentration in blood, cells and tissues of an organism (Liere et al., 1963). The brain is the most sensitive organ to the lack of oxygen. The condition of hypoxia causes disturbances of mental activities such as disturbances of memory, sight and speech, etc. Among the other factors, inhaling the air with lowered oxygen concentration can cause hypoxia. They can occur in tunnels during diving, flying with planes, climbing and staying on high mountains (Richalet et al., 1992; Obrenović, 2002; Obrenović and Nešić, 2000).

There are acute and chronic hypoxia conditions. Acute hypoxia implies the situation in which someone, whose natural environment is at low altitudes, is exposed to stay on high mountains. In the period of adaptation, we talk about the acute hypoxia. Under the chronic hypoxia, we understand the condition of a lowered amount of oxygen in the people whose natural environment is on high mountains. Some references (Obrenović, 2002; Obrenović and Nešić, 2000; Milivojević and Milivojević, 2010) show that the fundamental frequency of the speech signal changes due to hypoxia. This paper analyzes the effect of hypoxia on the fundamental frequency of speech as well as some dissonant spectrum ranges in relation to the change of the altitudes. Accordingly, the influence of dissonant spectrum ranges on the speech signal quality has been analyzed in the last decade. According to Joen et al. (2003), Kang (2004) and Milivojević and Balanesković (2009), it has been shown that it is possible to improve the quality of the speech signal in conditions of inconvenient ambient noise, as well as the quality of husky voice by filtering dissonant frequencies (Mokhtar et al., 2011; Wang et al., 2011). The algorithm for the quality improvement of the speech signal implies estimation of the fundamental frequency. On that basis, the algorithm defines and eliminates the dissonant spectral ranges. Furthermore, after the explanation of hypoxia and dissonant frequencies, the original algorithm for the calculation of dissonant coefficients' $k_{F\#}$, k_B and $k_{C\#}$ is

*Corresponding author. E-mail: dbrodic@tf.bor.ac.rs. Tel: +381 30 2496 772. Fax: +381 30 421 078.

given. From the testing results, the relation between the degree of hypoxia and dissonant coefficients will be set.

HYPOXIA

Hypoxia is a condition of insufficient concentration of oxygen in blood, cells and tissues, which causes functional disorder of organs, the nervous system and cells. Hypoxia can affect some organs, as well as the whole organism because insufficient concentration of oxygen cells die off, tissues decay or the function of many organs is disturbed: brain, lungs, heart, blood vessels, liver, and kidneys. Among all the organs in organism, the brain is the most sensitive to the insufficiency of oxygen. Hence, studying of this condition is of major importance. Hypoxia can occur due to:

1) Deficiency of oxygen in the atmosphere such as staying on high mountains.
2) Incidents in mines in underground pits, in aviation, cosmic flights, underwater activities, etc.
3) Lung diseases.
4) Disorder of the breathing center.
5) Diseases of the blood vessels.
6) Increased need of tissues for oxygen during extreme work of mussels, which usually happen to sportsmen and physical workers.

The hypoxia can cause the following effects on organism:

1) Lowering of mental activities (indicated through short memory, forgetfulness, slow thinking. sleepiness, euphoria, headache, nausea, sight and speech disturbances and finally jerks, convulsions and coma).
2) Lowering of the working capability of muscles (manifested in slow walk, feeling of powerlessness, weak and slow reflexes, bad coordination of motor movements and bad accommodation of eyes).
3) Depression of the respiratory center (losing consciousness, coma and death).

The hypoxia affects the changing of total functional state of an organism. Hence, it is a powerful stress generic factor that affects the human speech apparatus as well. Some investigations have been carried out about the effect of hypoxia on the speech as a consequence of changing the altitudes (Obrenović, 2002; Milivojević and Milivojević, 2010). The concentration of oxygen in the atmospheric air is approximately 21% when the pressure is 1 bar at the altitudes. With higher altitudes, one gets less oxygen, that is, hypoxia phenomena intensifies. Investigations have shown that at 1600 m altitudes, a considerable effect of hypoxia appears. Hence, these altitudes are called the reaction threshold. The highest altitudes where human settlements were formed and survived are 5500 m altitudes (the Andes, Tibet, SheSharpLase Himalaya) (Obrenović and Nešić, 2000). This height is considered to be the outmost limit of human adaptability. A healthy man who is not accustomed to this height, can preserve his full working ability and his state of consciousness for about 30 min at that height. However, a man can adapt to this height and to stay there for a long time. On the other hand, examinations have shown that on the height of 6700 m, it is impossible to survive for a long time. Without any adequate protective equipment, a man can survive for about 10 min at that height.

Physical characteristics of the speech signal under the influence of acute hypoxia have been analyzed according to Obrenović (2002). The analysis of the speech fundamental frequency, in two test groups has been made. Investigation was carried out on the people who live naturally at 400 and 1600 m altitudes. The analyses have been carried out to examine the effect of acute and chronic hypoxia. The acute hypoxia effect was analyzed by the measurements of the fundamental frequency on the tested persons who live at 400 and 1600 m altitudes. The chronic hypoxia was analyzed on the tested persons who live at 1600 m altitudes. The examinations were performed in cases of separately uttered vowels: A, E, I, O and U. It was stated that, among others, it comes to slight increasing of the fundamental frequency due to the acute hypoxia. The effect of the chronic hypoxia led to considerable increase of the fundamental frequency, that is, from 140 to 170 Hz.

So far, the effect of hypoxia to dissonant spectral ranges of the speech signal has not been analyzed. In this paper, the results of the analysis of energy distribution in dissonant spectral ranges have been presented. To determine the limits of the dissonant spectral ranges of the speech signal, the musicological definition of the dissonant ranges and the realized relation to the spectrum of the speech signal has been described.

MUSICOLOGICAL DEFINITION OF THE DISSONANT FREQUENCIES

The theory of music defines the fundamental features of the sound by:

1) Duration
2) Intensity
3) Color.

The expression color applies to the sound in a metaphorical way, which points out to the complexity of this feature of the sound. The source of a sound generates a sound with the fundamental frequency which represents the primary tone. Accordingly, it generates the overtones which aliquot in relation to the primary tone. Different number of the present aliquot (lat. *aliquot* - several times) and their various relative intensity within the total sounding determine the color of the sound (Joen et al., 2003; Kang, 2004).

The frequency of the musically defined tones in relation to the primary tone in an interval of one octave is determined by:

$$F_k = F_0 \times 2^{\left(k/12\right)}, \quad k = 0,1,...,12 \tag{1}$$

where F_0 is the frequency of the primary tone and F_k the frequency of the k-th halftone. Octave consists of 12 halftones. In relation to the primary tone, the halftones form intervals. An interval is defined by the relation of the frequency of a halftone and the frequency of the primary tone. Fractions F_k/F_0, for k = 0,1,...,12, which present individual intervals (1/1, 135/128, 9/8, 6/5, 5/4, 4/3, 45/32, 3/2, 8/5, 27/16, 9/5, 15/8, 2/1) present the approximation of the real value in Equation 1. Interval classification according to their sounding is realized based on the fraction which it describes. If the fraction is simpler, then the interval as an accord of tones is more stable that is, more consonant. If the fraction is more complex, then the stability of the interval is smaller. Hence, the dissonance is greater.

Consonance and dissonance are not sharply delimited. However, they make together one differentiated scale, from total stability on one end of the scale to total instability on the other end. Within the scale, we distinguish the following:

1) Perfect (complete) consonances [prima (1/1), octave (2/1), quinta (3/2) and quarta (4/3)]
2) Imperfect (incomplete) consonances [big tierce (5/4), big sixth (5/3), small tierce (6/5) and small sixth (8/5)]
3) Imperfect (incomplete) dissonances [small seventh (9/5) and big second (9/8)]
4) Perfect (complete) dissonances [small second (135/138), three tones or excessive quarta (45/32) and big seventh (15/8)].

From the view-point of experience, that is, perception of the sound; if the sound is pleasant or restful, then the musical interval is defined as consonant. However, if the sound is unpleasant or rough

(Milivojević and Balanesković, 2009), the musical interval is a dissonant.

Half tones' frequencies, which together with the primary tone make consonances in all octaves within the audible range (F_d - dissonant halftone frequency), are defined as:

$$F_d = F_0 \times 2^{\left(n+k/12\right)}, \quad n = 0,1,...,7;\ k = 1,6,11 \tag{2}$$

where F_0 is the frequency of the primary tone, n is the number of the octaves and k is the number of dissonant halftones in individual octaves. Considering the tone C as the referent one, that is, as the primary tone, then its half tones are F# (k = 6), B (k = 11) and C# (k = 1) as well as their harmonics in all the octaves.

Digital processing of the speech signal in order to determine the emotional state of the speaker automatically (happiness, sadness, anger, worry, boredom, ...) implies the analysis of the fundamental frequency and its relation to other frequencies in one octave, too (Yang and Lugger, 2010; Murray et al., 1996; Ruiz et al., 1996).

ALGORITHM FOR THE ESTIMATION OF DISSONANT COEFFICIENTS

The aim of the proposed algorithm is determination of the dissonant energy in the spectrum depending on hypoxia caused by the change of the altitudes. The proposed algorithm consists of the following steps:

Step 1: The speech signal $x(n)$, where n = 1,...,L is being divided into frames whose length is N when the window function $w(n)$ is applied with overlapping of $N/4$ frames.

Step 2: For every frame, the fundamental frequency F_0 is being determined (Pang et al., 2000; Milivojević et al., 2006; Milivojević and Mirković, 2009).

Step 3: The energy in the range of the fundamental frequency harmonic is being determined.

Step 4: The energy in dissonant frequencies' ranges which correspond to halftones' F#, B and C# are being determined.

Step 5: The dissonant coefficients' $k_{F\#}$, k_B and $k_{C\#}$ are being determined.

The dissonant coefficients' $k_{F\#}$, k_B and $k_{C\#}$ point to the percentage of the energy participation in dissonant spectral range F#, B and C# in relation to the energy around the fundamental frequency and its harmonics. The lower value of dissonant coefficients represents the increase of the dissonant energy.

The realization of the Steps 3 to 5 is as follows:

Step 3 of the proposed algorithm is realized in the following way:
For n=1 to 20 do:

a) Determination of the limits for the range whose center is the n-th harmonic of the fundamental frequency:

$$k_{F0\_bl} = \left(nF_0 - \frac{w_0}{2}\right)\frac{F_S}{NFFT}, \tag{3}$$

$$k_{F0\_br} = \left(nF_0 + \frac{w_0}{2}\right)\frac{F_S}{NFFT}. \tag{4}$$

b) Determination of the energy in the range of the n-th harmonic of the fundamental frequency:

$$E_{F_0} = \sum_{k=K_{F0\_bl}}^{K_{F0\_br}} S^2(k) \tag{5}$$

where F_0 is the fundamental frequency, F_S is the sampling frequency, S = abs($FFT(x,NFFT)$), $NFFT$ is the length FFT, w_0 is the width of the range of one halftone, k_{F0_bl} and k_{F0_br} are the limits of spectral components, respectively.

Step 4 of the proposed algorithm is realized in the following way:
For n=1 to 7 do:

a) Determination of the harmonics for the dissonant frequency:

$$F_{dF\#} = F_0 \times 2^{\left(n+6/12\right)} \tag{6}$$

$$F_{dB} = F_0 \times 2^{\left(n+11/12\right)} \tag{7}$$

$$F_{dC\#} = F_0 \times 2^{\left(n+1/12\right)} \tag{8}$$

b) Determination of the limits for the range whose center is the n-th harmonic of the dissonant frequency:

(b.1) The limit of F# range:

$$k_{F\#\_bl} = \left(nF_{dF\#} - \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{9}$$

$$k_{F\#\_br} = \left(nF_{dF\#} + \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{10}$$

(b.2) The limit of B range:

$$k_{B\_bl} = \left(nF_{dB} - \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{11}$$

$$k_{B\_br} = \left(nF_{dB} + \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{12}$$

(b.3) The limit of C# range:

$$k_{C\#\_bl} = \left(nF_{dC\#} - \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{13}$$

$$k_{C\#\_br} = \left(nF_{dC\#} + \frac{w_0}{2}\right)\frac{F_S}{NFFT} \tag{14}$$

(c) Determination of energy in the range of the n-th dissonant harmonic:

$$E_{F\#} = \sum_{k=K_{F\#\_bl}}^{K_{F\#\_br}} S^2(k) \tag{15}$$

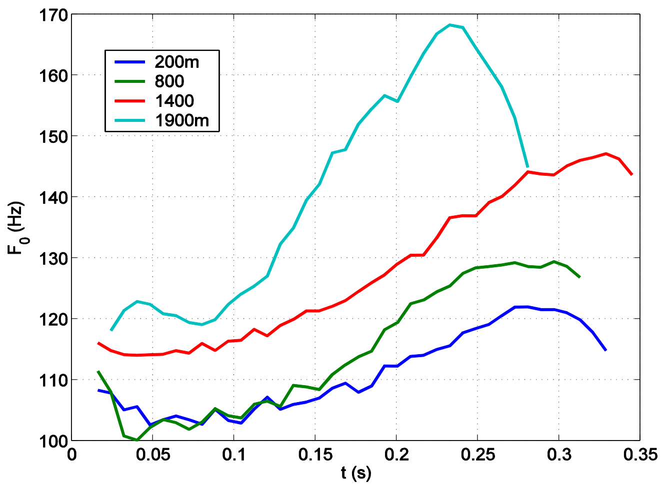

Figure 1. Trajectories of the fundamental frequency for 200, 800, 1400 and 1900 m altitudes (vowel A).

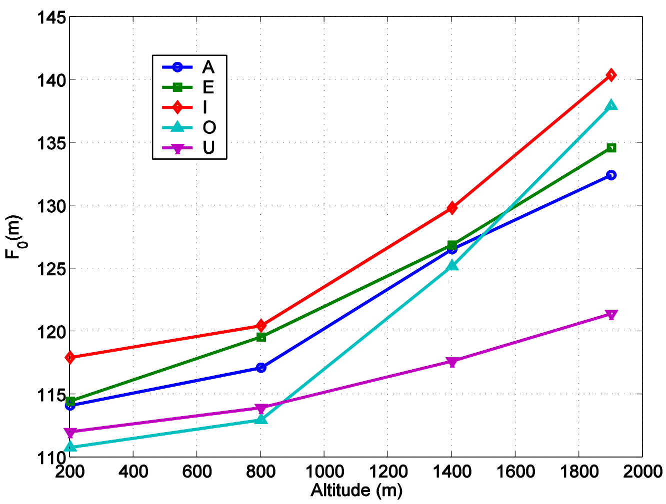

Figure 2. Mean value of the fundamental frequency for vowels in function of the altitudes.

$$E_B = \sum_{k=K_{B\#\_bl}}^{K_{B\#\_br}} S^2(k) \quad (16)$$

$$E_{C\#} = \sum_{k=K_{C\#\_bl}}^{K_{C\#\_br}} S^2(k) \quad (17)$$

end

Step 5 The dissonant coefficients are:

$$k_{F\#} = \frac{E_{F0}}{E_{F\#}} \quad (18)$$

$$k_B = \frac{E_{F0}}{E_B} \quad (19)$$

$$k_{C\#} = \frac{E_{F0}}{E_{C\#}} \quad (20)$$

EXPERIMENTAL RESULTS AND ANALYSIS

The speech database

For the efficiency testing of the proposed algorithm for estimation of dissonant coefficients, a test group was formed. It is composed of the persons whose natural environment was at 200 m altitudes (the city of Niš, Serbia). Test group was made of five persons. Measurements were performed at 200, 800, 1400 and 1900 m altitudes above the sea level. Each person articulated the vowels A, E, I, O and U three times on each height with a pause of 5 min between the utterances. The speech signal was stored in the form of *wav* file and thus, the database of the speech signals (F_S=16 kHz) was formed (Milivojević and Milivojević, 2010).

Results

The described algorithm for processing speech signal was applied with Hamming window function, N = 512 and *NFFT* = 2048 over the speech signals database. Trajectories of the fundamental frequency for different heights are shown in Figure 1. The mean value of the fundamental frequency for vowels as a function of the altitudes is shown in Figure 2. The mean value of dissonant coefficients for vowels A, E, I, O and U is shown in: a) $k_{F\#}$ (Figure 3), b) k_B (Figure 4), and c) $k_{C\#}$ (Figure 5). The mean value of dissonant coefficients for all vowels is shown in Figure 6.

Analysis of the results

According to the analysis of the results shown in Figures 2 to 6 and the results from Milivojević and Milivojević (2010), the following conclusions are drawn:

1) With the increase of the altitudes, the fundamental frequency increases (in relation to 200 m altitudes, it increases for 2.58% at 800 m altitudes, 9.95% at 1400 m altitudes and 17.1% at 1900 m altitudes), which agrees considerably with the results from Obrenović (2002) and Milivojević and Milivojević (2010).
2) With the increase of the altitudes, the value of the dissonant coefficient $k_{F\#}$ increases also (in relation to 200 m altitudes, it increases for 13.76% at 800 m altitudes,

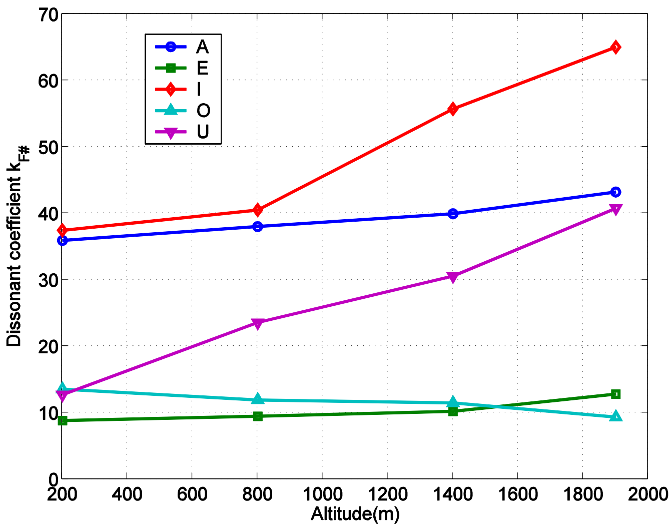

Figure 3. Mean value $k_{F\#}$ for vowels in function of the altitudes.

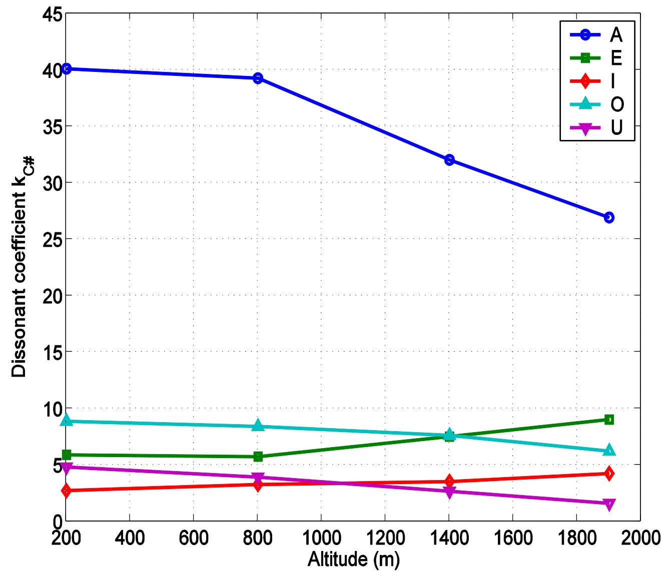

Figure 5. Mean value $k_{C\#}$ for vowels in function of the altitudes.

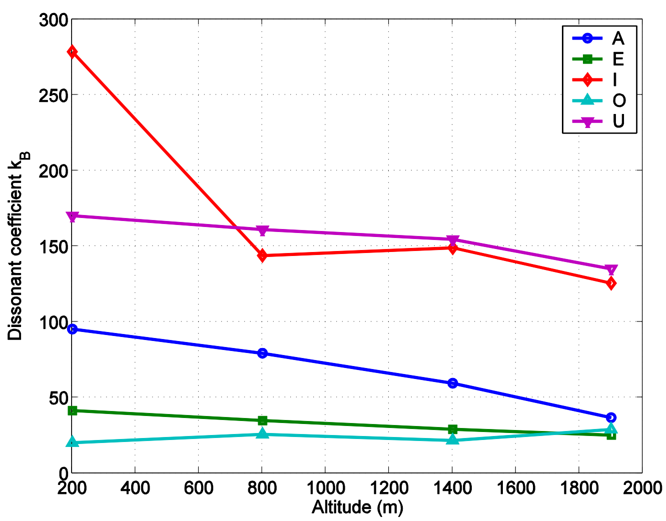

Figure 4. Mean value k_B for vowels in function of the altitudes.

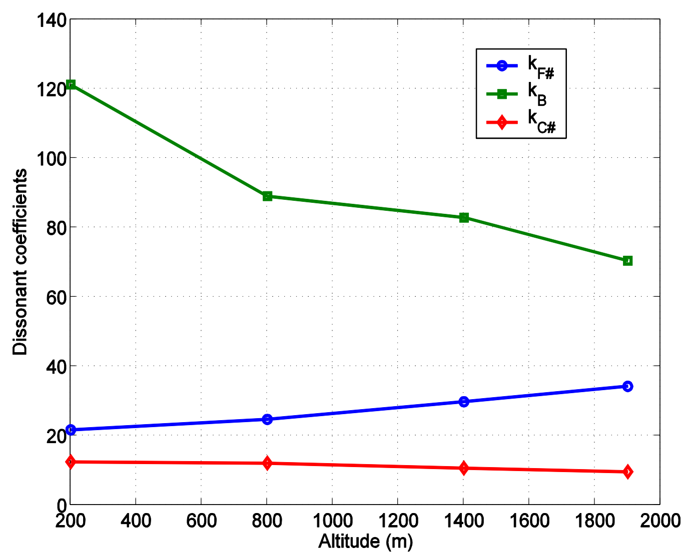

Figure 6. Mean value of dissonant coefficients $k_{F\#}$, k_B and $k_{C\#}$ for all vowels.

36.69% at 1400 m altitudes and 57.76% at 1900 m altitudes), that is, the energy in dissonant F# ranges is reducing (Milivojević and Milivojević, 2010).

3) With the increase of the altitudes, the value of the dissonant coefficient k_B decreases (in relation to 200 m altitudes, it decreases for 17.1% at 800 m altitudes, 26.53% at 1400 m altitudes and 41.87% at 1900 m altitudes), that is, the energy in dissonant B ranges is increasing.

4) With the increase of the altitudes the value of the dissonant coefficient $k_{C\#}$ decreases (in relation to 200 m altitudes it decreases for 2.87% at 400 m altitudes, 14.51% at 1400 m altitudes and 23.01% at 1900 m altitudes), that is., the energy in dissonant C# ranges is increeasing. According to the analysis of dissonant coefficients' $k_{F\#}$, k_B and $k_{C\#}$, it is possible to recap about the degree of hypoxia. Such data can be used in many situations. For instance, during the flight on great heights, the pilot's mask can fail, or it can come to the decompression of the plane. In such cases, it is possible on the basis of processing the pilot's conversation with the flight control, to discover signs of hypoxia in its initial stage and take adequate measures in order to prevent catastrophic consequences (Obrenović and Nešić, 2000).

CONCLUSION

In this paper, the speech signal has been analyzed in order to determine the degree of hypoxia. Accordingly, a new method which is based on the analysis of dissonant F#, B and C# spectral range was proposed. The coefficients' $k_{F\#}$, k_B and $k_{C\#}$ was calculated for vowels A, E, I, O and U, articulated by the group of tested persons at different altitudes: 200, 800, 1400 and 1900 m. By the analysis of dissonant coefficient $k_{F\#}$, k_B, and $k_{C\#}$, it is possible to use them as indicators of hypoxia, and even to determine the degree of hypoxia.

REFERENCES

Joen B, Kang S, Baek SJ, Subg KM (2003). Filtering of a dissonant frequency based on improved fundamental frequency estimation for speech enhancement. IEICE Trans. Fundamentals, 86(8): 2063-2064

Kang S (2004). Dissonant frequency filtering technique for improving perceptual quality of noisy speech and husky voice. Signal Processing, 84: 431-433.

Liere V, Edward J, Clifford J (1963). Hypoxia. The University of Chicago and London.

Milivojević M, Milivojević Z (2010). Dissonant f# interval of the speech signal as a hypoxia indicator. Inform. Tech. Žabljak, Montenegro.

Milivojević Z, Balanesković D (2009). Enhancement of the perceptive quality of the noisy speech signal by using of DFF-FBC algorithm. Facta Universitatis. S.El. Energ, 22(3): 379-392.

Milivojević Z, Mirković M (2009). Estimation of the fundamental frequency of the speech signal modeled by the SYMPES method. AEU – Int. J. of Elect. Comm. 63: 200-208.

Milivojević Z, Mirković M, Milivojević S (2006). An estimate of fundamental frequency using PCC interpolation – comparative analysis. Inform. Tech. and Control 35(2): 131-136.

Mokhtar N, Arof H, Mahamd-Adikan FR, Mubin M (2011). Real time noise-speech discrimination in time domain for speech recognition application. Scientific Research and Essays, 6(1): 18-22.

Murray I, Baber C, South A (1996). Towards a definition and working model of stress and its effects on speech. Speech Communication 20: 3-12.

Obrenović J (2002). Fundamental frequency of the voice in relation to hypoxia as a stressor. Facta Universitatis, S. Philosophy, Sociol. Psychol., 2(9): 683-689.

Obrenović J, Nešić M (2000). Cognitive values of speech signal under the conditions of acute hypoxia. Psychology. University of Niš 2(2): 127-137.

Pang HS, Baek SJ, Sung KM (2000). Improved fundamental frequency estimation using parametric cubic convolution. IEICE Trans. Fundamentals. E83-A 2747-2797.

Richalet JP, Kacimi R, Antezana AM (1992). The control of chronotropic function in hypobaric hypoxia. Int. J. Sports Med., 13(1): 22-24.

Ruiz R, Absil E, Harmegnies B, Legros C, Poch D (1996). Time and spectrum related variability in stressed speech under laboratory and real conditions. Speech Communication, 20: 111-129.

Yang B, Lugger M (2010). Emotion recognition from speech signals using New Harmony features. Signal Processing 90: 1415-1423.

Wang B, Li Z, Ye W (2011). Speech control of a teleoperated mobile humanoid Robot. Int. J. Phys. Sci., 6(18): 4478-4491.