# The Effects of the Acute Hypoxia to the Fundamental Frequency of the Speech Signal

Zoran N. MILIVOJEVIĆ[1], Marina MILIVOJEVIĆ[2], Darko BRODIĆ[3]
[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 Bor, Bor, V.J. 12, 19210 Bor, Serbia*
[1,2] *zoran.milivojevic@jotel.co.rs;* [3] *dbrodic@tf.bor.ac.rs*

***Abstract*—When people that live at the small altitudes (up to 400 m above the sea level) climb on the mountain, they are exposed to the effects of an acute hypoxia. As a consequence, theirs oxygen concentration decreases in the tissue. This paper presents the analysis of the acute hypoxia effects to the speech signal at the altitudes up to 2600 m above the sea level. For the experiment, the articulation of vowels (A, E, I, O, U) from the test group of persons was recorded at different altitudes, which creates the speech signal database. The speech signal from database is processed by the original algorithm. As the results, the fundamental frequency and the energy of dissonant intervals of speech signal are obtained. Furthermore, the acute hypoxia effect to the energy distribution in the dissonant intervals of the speech signal is analyzed. At the end, the comparative analysis of the acute hypoxia effects shows that the level of the hypoxia can be determined by the change of the fundamental frequency and the energy of the dissonant intervals of speech signal. Hence, it is possible to bring conclusions about the degree of hypoxia, which in many situations can be of importance for avoiding catastrophic consequences.**

***Index Terms*—speech analysis, speech processing, fundamental frequency, dissonant frequency, acute hypoxia.**

## I. Introduction

Hypoxia represents a condition of decreased concentration of oxygen in blood, cells and tissues of an organism [1]. The brain is the most sensitive organ, so the condition of hypoxia causes disturbances of mental activities (disturbances of memory, sight and speech, etc.). Among other factors which can cause hypoxia there is inhaling the air with lowered oxygen concentration. This circumstance can occur in tunnels, during diving, during flights with planes, and during stay on high mountains [2-4].

There are acute and chronic hypoxia conditions. Under acute hypoxia we understand the situation when someone whose natural environment is at low altitudes, is exposed the stay on high mountains. In the period of adaptation we talk about the acute hypoxia. Under the chronic hypoxia we understand the condition of lowered amount of oxygen in people whose natural environment is on high mountains. In [3-4] it is shown that the fundamental frequency of the speech signal changes due to hypoxia.

According to [5], the processing of the fundamental frequency as well as its relation to the other frequencies in spectrum can designate the emotional state of the speaker (happiness, sadness, anger, anxiety, boredom, disgust, and neutral) and the stress [6-7]. However, the major unknown is the impact of the acute hypoxia to the energy in dissonant spectral ranges. If there is a link, then degree of the hypoxia can be determined. The dissonant ranges known as devils intervals are defined by the theory of music [8-9].

In this paper, the hypoxia and dissonant frequencies are explained. Furthermore, the original algorithm for the calculation of fundamental frequency and dissonant coefficients $k_{F\#}$, $k_{B}$ and $k_{C\#}$ is given. This algorithm is applied to the speech database. The speech database is formed from the voices of speakers, which live at the small altitudes (up to 400 m above the sea level). These speakers pronounce the vowels: A, E, I, O, and U at the different altitudes (up to 2600 m above the sea level). Speech is recorded on PC. It is archived in the WAV format files on hard disk. From the obtained results the relation between the degree of hypoxia and dissonant coefficients is set.

The organization of the paper is the following. In Section II hypoxia is described. In Section III definition of dissonant frequencies is presented. In Section IV a new algorithm for calculation of dissonant coefficients is described. In Section V the testing results and the comparative analysis are presented. Section VI makes conclusion.

## II. Hypoxia

Hypoxia is a condition of insufficient concentration of oxygen in blood, cells and tissues. It causes functional disorder of organs, the nervous system and cells. Due to 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. Consequently, hypoxia can affect some organs, as well as the whole organism. The brain is the most sensitive to lowering and insufficiency of oxygen. Hence, studying of this condition is in the focus. The causes for hypoxia can be multiple: deficiency of oxygen in the atmosphere (staying on high mountains, during incidents in mines in underground pits, in aviation, cosmic flights, underwater activities etc.), lung diseases, disorder of the breathing center, diseases of the blood vessels, increased need of tissues for oxygen during extreme work of mussels which usually happen to sportsmen and physical workers.

Characteristics of the effect of hypoxia on the organism are: a) 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), b) lowering of the working capability of muscles (manifested in slow walk, feeling of powerlessness, weak and slow reflexes, bad coordination of motor movements, bad accommodation of eyes), and c)

Digital Object Identifier 10.4316/AECE.2012.02010

depression of the respiratory center (losing consciousness, coma and death). Hypoxia affects the changing of total functional state of an organism as well as the human speech apparatus itself. Some investigations have been carried out concerning the effect of hypoxia on the speech as a consequence of changing the altitudes [3, 5]. Concentration of oxygen in the atmospheric air is approximately 21% when the pressure is 1 bar at the altitudes. With higher altitudes the air density decreases and by that the quantity of oxygen. With higher altitudes one gets less oxygen, i.e. hypoxia intensifies. Investigations have shown that at 1600 m altitudes a considerable effect of hypoxia appears. For that reason these altitudes are called the reaction threshold. The highest altitudes where human settlements were formed and survived are 5500 m altitudes [4]. This height is considered to be the outmost limit of human adaptability power. A healthy man, who is not accustomed to this height, can preserve his full working ability and state of full consciousness for about 30 minutes. 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 not possible to survive for a long time. Without any adequate protective devices a man can survive for at most 10 minutes.

Physical characteristics of the speech signal under the influence of acute hypoxia have been analyzed in [3]. The results of the fundamental frequency analysis in two test groups have been presented, by those who live naturally at 400 m altitudes and by those who live at 1600 m altitudes. The analyses have been made in order to investigate the effect of acute and chronic hypoxia. The effect of the acute hypoxia has been analyzed due to measurements of the fundamental frequency on the tested persons who live at 400 m altitudes, and those who live at 1600 m altitudes. Furthermore, the chronic hypoxia has been analyzed on the tested persons who live at 1600 m altitudes. The examinations have been performed in cases of separately uttered vowels A, E, I, O and U. It was stated that it comes to slight increase of the fundamental frequency because of the acute hypoxia. Consequently, the effect of the chronic hypoxia led to considerable increase of the fundamental frequency, i.e. from 140 to 170 Hz. However, the effect of hypoxia to dissonant spectral ranges of the speech signal has not been analyzed in the literature so far. Hence, in this paper the results of the analysis of energy distribution in dissonant spectral ranges have been presented. In order to determine the limits of dissonant spectral ranges of the speech signal, the musicological definition of dissonant ranges and the realized relation to the spectrum of the speech signal has been described in the next chapter.

## III. MUSICOLOGIC DEFINITION OF DISSONANT FREQUENCIES

The theory of music defines the fundamental features of the sound: a) duration, b) intensity, and c) 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 (the primary tone) as well as the overtones (aliquoties in relation to the primary tone). Different number of the present aliquoties and their relative intensity within the total sounding, determine the color of a sound [9].

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. In relation to the primary tone half tones frequencies, which together with the primary tone make consonances in all octaves within the audible range, are defined as:

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

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

## IV. ALGORITHM FOR THE ESTIMATION OF DISSONANT COEFFICIENT

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 $x(n)$, $n$ = 1, ..., $L$ is being divided into frames. Theirs length is $N$, when the window function $w(n)$ with overlapping of frames $N/4$ is applied.

**Step 2**: For every frame the fundamental frequency $F_0$ is being determined [10-12].

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

**Step 4**: The energy in dissonant $F\#$, $B$ and $C\#$ ranges is being determined.

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

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

**Step 3** is realized in the following way:

**FOR** n = 1 **TO** 20 **DO**

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

$$k_{F_0\_bl} = \left(nF_0 - \frac{w_0}{2}\right)\frac{F_S}{NFFT}\,, \tag{3}$$

$$k_{F_0\_br} = \left(nF_0 + \frac{w_0}{2}\right)\frac{F_S}{NFFT}\,, \tag{4}$$

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

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

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.

**END**

**Step 4** is realized in the following way:

**FOR** n = 1 **TO** 7 **DO**

**(a)** Determination of the dissonant frequency harmonics:

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

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

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

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

**(b.I)** The limits of F# range are:

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

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

**(b.II)** The limits of $B$ range are:

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

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

**(b.III)** The limits of C# range are:

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

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

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

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

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

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

**END**

**Step 5** Dissonant coefficients are:

$$k_{F\#} = E_{F_0} / E_{F\#}, \tag{18}$$

$$k_B = E_{F_0} / E_B, \tag{19}$$

$$k_{C\#} = E_{F_0} / E_{C\#}. \tag{20}$$

## V. Experimental Results and Analysis

### A. The speech database

For testing the efficiency of the proposed algorithm for estimation of dissonant coefficients, a test group was formed composed of persons whose natural environment was at 200 m altitudes (Niš, Serbia). That test group was made of 5 persons (male from 18 to 50 years old). Further, the measurements were performed at 200, 800, 1400, 1800, 2200, and 2600 m altitudes (mountain Pelister, Macedonia). Every tested person articulated the vowels A, E, I, O and U three times on each height with a pause of 5 minutes between the utterances. The speech signal was stored on the hard disc in the form of wav file and thus the speech signals database ($F_S$=16 kHz) was formed.

### B. Results

The algorithm was applied to the speech signals database with the following parameters: a) Hamming window function, b) $N = 512$, and c) $NFFT = 2048$.

The results are shown in Figures 1-6. Figure 1 presents the trajectories of the fundamental frequency for different persons at different heights. The mean value of the fundamental frequency for vowels in the function of the altitudes is shown in Figure 2. The mean value of dissonant coefficients for vowels A, E, I, O and U is presented in the following figures: 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 presented in Figure 6.

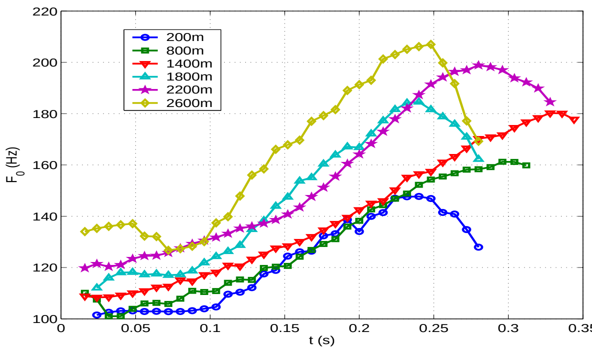


Figure 1. Trajectories of the fundamental frequency for 200, 800, 1400, 1800, 2200 and 2600 m altitudes (vowel A).

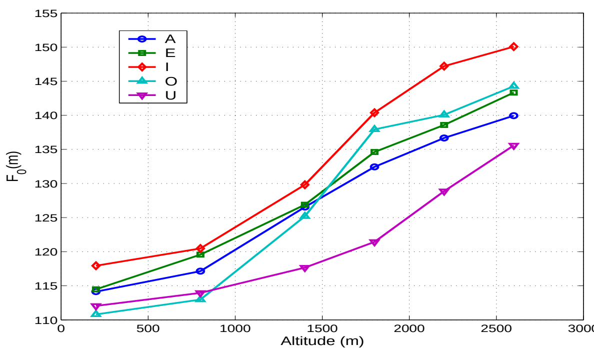


Figure 2. Mean value of the fundamental frequency for vowels in function of the altitudes.

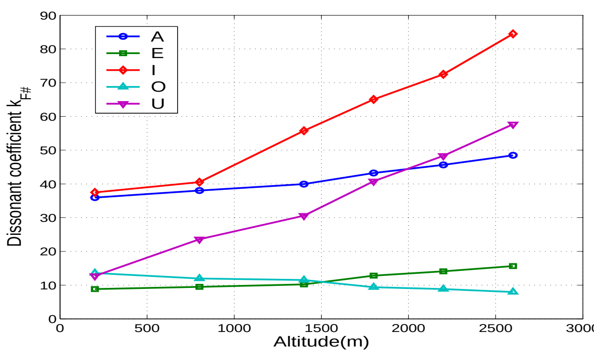


Figure 3. Mean value $k_{F\#}$ for vowels in function of the altitudes.

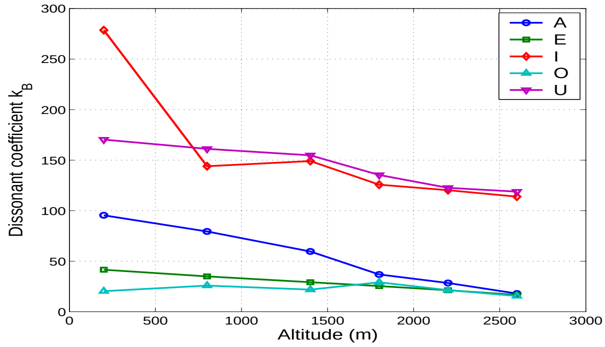

Figure 4. Mean value $k_B$ for vowels in function of the altitudes.

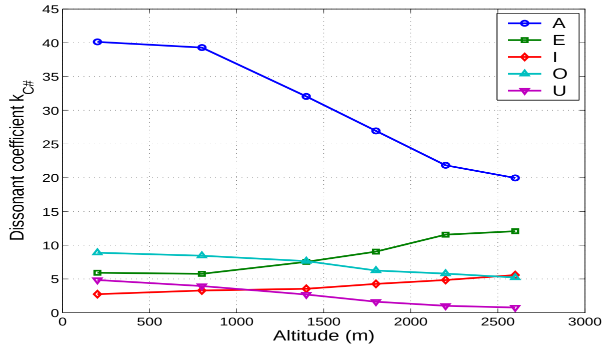

Figure 5. Mean value $k_{C\#}$ for vowels in function of the altitudes.

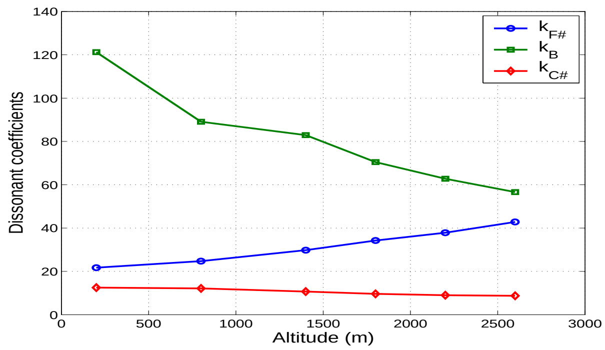

Figure 6. Mean value of dissonant coefficients $k_{F\#}$, $k_B$ and $k_{C\#}$ for all vowels.

### *C. Analysis of the results*

According to the analysis of the results presented in Figures 1-5 and the results from [5] the following conclusion can be derived:

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, 17.1% at 1800 m altitudes, 21.42% at 2200 m altitudes and 25.25% at 2600 m altitudes. These values agreed with the results given in [3].

2. With the increase of the altitudes the value of the dissonant coefficient $k_{F\#}$ increases. In relation to 200 m altitudes, it increases for 13.76% at 800 m altitudes, 36.69% at 1400 m altitudes, 57.76% at 1800 m altitudes 74.42% at 2200 m altitudes and 97.34% at 2600 m altitudes. Hence, it comes to lowering of energy in dissonant F# ranges.

3. With the increase of the altitudes the value of the dissonant coefficient $k_B$ decreases. In relation to 200m altitudes, it decreases for 17.1% at 800 m altitudes, 26.53% at 1400m altitudes, 41.87% at 1800 m altitudes, 48.21% at 2200 m altitudes and 53.31% at 2600 m altitudes. Consequently, it comes to rising of energy in dissonant B ranges.

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, 23.01% at 1800 m altitudes, 27.94% at 2200 m altitudes and 30.24% at 2600 m altitudes. It comes to raising of energy in dissonant C# ranges.

According to the analysis of dissonant coefficients $k_{F\#}$, $k_B$ and $k_{C\#}$, it is possible to recapitulate 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 base of processing of 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 [3-4].

## VI. Conclusion

In order to determine the degree of hypoxia the speech signal has been analyzed. In the paper, a new method which is based on the analysis of dissonant F#, B and C# spectral ranges are proposed. The coefficients $k_{F\#}$, $k_B$ and $k_{C\#}$ were calculated for vowels A, E, I, O and U, articulated by the group of tested persons at different altitudes (200, 800, 1400, 1800, 2200 and 2600 m). On the base of the analysis of $k_{F\#}$, $k_B$ and $k_{C\#}$ of dissonant coefficient, it is possible to use them as indicators of hypoxia, and even for determination of the degree of hypoxia.

## References

[1] V. Liere, J. Edward, J. Clifford, Hypoxia. The University of Chicago and London, 1963.
[2] Richalet J.P., R. Kacimi, A.M. Antezana, "The control of chronotropic function in hypobaric hypoxia," Int. J. Sports Med. 13:S22-S24, 1992.
[3] J. Obrenovic, "Fundamental frequency of the voice in relation to hypoxia as a stressor," Facta Universitatis, Series: Philosophy, Sociology and Psyhology, Vol. 2, No. 9, pp. 683-689, 2002.
[4] J. Obrenović, M. Nešić, Cogntive values of speech signal under the conditions of acute hypoxia, Psihologiya, University of Nis, Vol. 2, No. 2, pp. 127-137, 2003.
[5] B. Yang , M. Lugger, "Emotion recognition from speech signals using new harmony features", Signal Processing, Vol. 90, pp. 1415–1423, 2010.
[6] I. Murray, C. Baber, A. South, "Towards a definition and working model of stress and its effects on speech", Speech Communication, Vol. 20, pp. 3-12, 1996.
[7] R. Ruiz, E. Absil, B. Harmegnies, C. Legros, D. Poch, "Time and spectrum related variabilities in stressed speech under laboratory and real conditions", Speech Communication, Vol. 20, pp. 111-129, 1996.
[8] S. Kang, "Dissonant frequency filtering technique for improving perceptual quality of noisy speech and husky voice", Signal Processing, Vol. 84, pp. 431-433, 2004.
[9] Z. Milivojević, D. Balanesković, "Enhancement Of The Perceptive Quality Of The Noisy Speech Signal By Using Of DFF-FBC Algorithm", Facta Universitatis, Ser.: Elec. Energ. Vol. 22, No. 3, pp. 379-392, December 2009.
[10] H.S. Pang, S.J. Baek, Sung KM, "Improved Fundamental Frequency Estimation Using Parametric Cubic Convolution", IEICE Trans. Fundamentals, E83-A, pp. 2747-50, 2000.
[11] Z. Milivojevic, M. Mirkovic, S. Milivojevic, "Fundamental Frequency Estimation of the Speech Signal Compressed by MP3 Algorithm Using PCC Interpolation", Advances in Electrical and Computer Engineering AECE, Vol. 10, No. 1, pp. 18-22, 2010.
[12] Z. Milivojevic, M. Mirkovic., "Estimation of the fundamental frequency of the speech signal modeled by the SYMPES method", AEU - International Journal of Electronics and Communications, Vol. 63, pp. 200-208, 2009.