The Effect of Long-Term Use of Drugs on Speaker's Fundamental Frequency

Andrey Raev [1], Yuri Matveev [1], Tatiana Goloshchapova [2]
[1] Speech Technology Center, St. Petersburg, RUSSIA
{raev, matveev}@speechpro.com
[2] Federal Drug Control Service of the Russian Federation, Moscow, RUSSIA

Abstract

In this paper we investigate speech recordings before and after speaker's drug-abuse treatment, and show that there are no statistically significant dependences in distortions of speaker's fundamental frequency on different groups of drugs and on the degree of drug intoxication. Changes in the fundamental frequency are not regular and have no general nature. The main reason for changing the fundamental frequency is a change of speaker emotional state, rather than drug addiction treatment. Exploring the effect of the duration of narcotic drug usage on the speaker's fundamental frequency showed that speakers with prolonged use of drugs of the heroin group have a tendency to decrease the fundamental frequency by about 3% per year.

1. Introduction

It is a widely accepted hypothesis that drug intoxication as other factors such as an emotional state (for example, stress) and a functional state (for example, illness) influence the way a person speaks and his vocal parameters. For example, in [1] were analyzed the general theoretical and applied researches on the impact of emotional and functional states of the person on the acoustic characteristics of his speech. It is shown that in most cases as the most informative acoustic correlates of emotional and functional states considered some of the frequency, timing and power characteristics of the voice signal. Noted that, as a rule, the sthenic state leads to an increase, and the asthenic - to a decrease of the fundamental and formant frequencies, and the intensity.

As noted in [2] a number of vocal parameters can be used as an indicators for the effects of psychoactive and antidepressant drugs, in particular fundamental frequency, spectral energy distribution, and pause length, changed with the patient mood state over different therapy periods. In [3] were analyzed following acoustic parameters: mean values and standard deviations of sound power, the fundamental frequency, frequencies of the first and second formants, and duration of utterances. After comparison of values of each emotional intonation with the corresponding parameters of the neutral intonation, it was found that the most pronounced in the percentage were deviations of the fundamental frequency and of the first formant. It was shown that the direction of these deviations do not depend on the semantic content of the utterance and its duration, speaker age and gender, although the absolute values of the frequencies of individual features of speakers influenced. In [4] was investigated the changes of the fundamental frequency and the shift of the formant frequencies for persons under different types of stress and emotions.

The problem of how drug intoxication influences on the speaker voice parameters was investigated in a relatively small number of studies, and there no known studies how these parameters are depend on different groups of drugs and the degree of drug intoxication, as well as the time of using of drugs. This is mainly due to the lack of needed speech databases to fulfill such studies. To fulfill the above studies, the authors collected a specialized speech database with samples of speech of speakers who are drug intoxicated (Speech Database Narco - SDBN). The SDBN was collected by STC Ltd (St. Petersburg, Russia) in a several specialized health care institutions on the request of the Federal Drug Control Service of the Russian Federation.

This paper presents the results of studies of the effect on the fundamental frequency of drug intoxication using the collected database. The results of these studies are important for expert

practice, and should be considered when developing automatic speaker identification systems [5], [6] and [7].

2. Description of the database of drug intoxicated speech

2.1. Collection of the database

All enrolled speakers in the SDBN have the personal number and the record containing the following information:

- Speaker's gender;
- Estimated degree of drug intoxication;
- Hypothetical type of the drug substance used.

We recorded only those speakers who were in drug intoxication. Detection of the drugged state and the degree of drug intoxication was carried out by narcologists (therapists for drug abusers at hospitals, experts at the out-patient department for drug addicts). The duration of the utterances is not less than 32 sec. Each speaker was recorded either once (in the drug intoxicated state) or twice (first time drug intoxicated, second time - after leaving the state of drug intoxication). Speaker records were made with a digital recorder at a distance of 1-2 meters from the speaker's lips, under the supervision of specially trained operators, in accordance with the instructions. When finished the recording, the recorded speech material is fully tapped and segmented with an audio editor program. The operator's speech, pauses and acoustic noises are deleted from the records.

2.2. The main characteristics of the collected database

The present study is based on the SDBN which contains the speech of 156 Russian speakers of both genders (67 female and 89 male speakers), recorded in drug intoxicated and sober conditions. The aim was to analyze the effect of drug intoxication on speaker's F0 with regard to the gender and the length of time speaker has been using drugs.

Aside from the speech signal we collected a number of metadata about speakers to allow statistical cross-testing for other factors than the main factor sober/intoxicated. In Tables 1-2 these metadata are summarized for speakers in the SDBN. Primarily, the speakers in the SDBN have different degrees of intoxication (see Table 1): mild, moderate, severe. The degree of drug intoxication was determined by narcologists. Secondary, the speakers are under the influence of different groups of drugs (see Table 1): plant-based (marijuana, hashish), semi-synthetic (heroin) and synthetic (amphetamine). In addition, in the SDBN presented speakers of different genders and different length of time of using drugs (see Table 2).

Table 1. Number of speakers with different degrees of intoxication and different groups of drugs

	Degree of drug intoxication			**Type of drugs used**		
	mild	moderate	severe	plant-based (marijuana, hashish)	semi-synthetic (heroin)	synthetic (amphetamine)
Number of speakers	39	64	52	10	142	5

Table 2. Number of speakers with different length of time of using drugs

	Length of time of using drugs, years				
	1	2	5	10	25
Number of speakers	66	21	24	34	11

3. Speech records analysis of drug using speakers

Based on the collected SDBN it was analyzed the effect of the degree of drug intoxication, different groups of drugs and the length of use of narcotic drugs on speaker's fundamental frequencies.

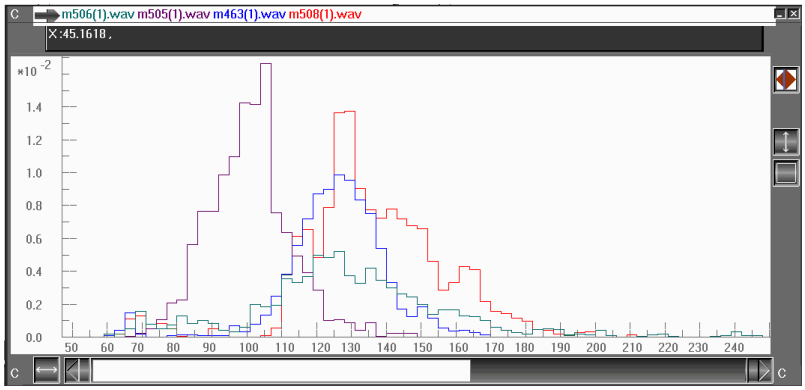
3.1. Features of the speaker in the state of drug intoxication

Analysis of the speaker's speech records before and after drug treatment indicates that the speaker's speech is changes slightly and irregularly.

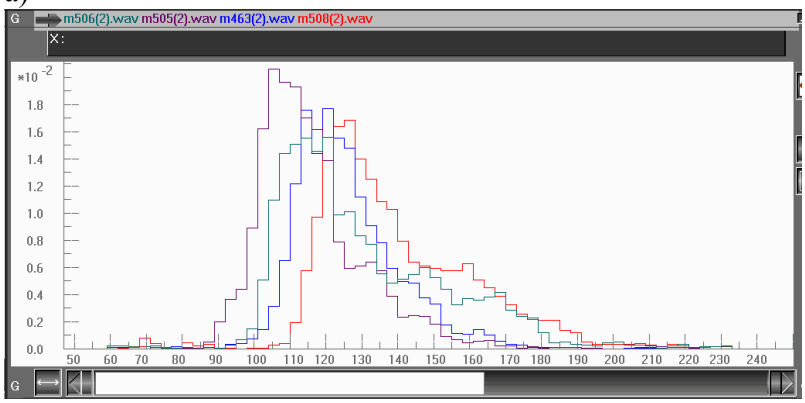
Fig. 1a and Fig. 1b shows the histograms of the fundamental frequencies (F0) of four men before and after drug treatment. After drug treatment, the F0 of one of the speakers is increased, of another is reduced, of other speakers is unchanged.

Fig. 1c shows the averaged histogram of F0 for the same four men before and after drug treatment. Parameters of the histogram have changed slightly: the mean and median of the distribution of the frequencies remained the same, and the variance decreased. But this was most likely due to changes in emotional state.

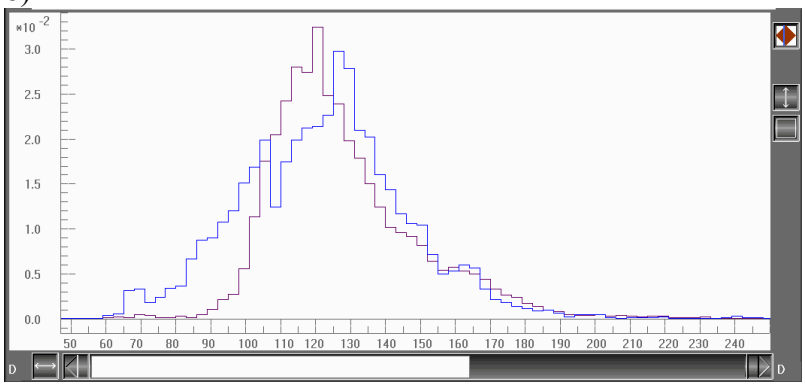
The situation is similar for women: the mode of the fundamental frequency distribution increased from 180 Hz to nearly 200 Hz, but the mean and median of the distribution is almost unchanged. Statistically significant regular changes are not observed. This suggests that the main reason for change of the fundamental frequency is a change in the emotional state of speakers, rather than drug addiction treatment.



a)



б)



c)

Fig. 1. Histograms of the fundamental frequencies of four men a) before drug treatment and b) after drug treatment, c) averaged histogram of the fundamental frequencies of men before (the right mode) and after (the left mode) drug treatment

3.2. The dependence of the speaker's fundamental frequency on different groups of drugs, the degree of drug intoxication and different length of time of drug use

There was carried out a statistical study of the effect on the fundamental frequency of the following characteristics:

- The degree of drug intoxication.

- The type of the drug substance used.
- The length of time of using drugs.

Considering approval of narcologists, that methods for determining the degree of drug intoxication are not currently available, we used a Mean Opinion Score (MOS) for the degree of intoxication and for the length of time of using drugs.

To determine the degree of speaker intoxication we primarily focused on the following symptoms: speech disorders (slurred speech), inability to understand what a doctor asks him to do, omission or misunderstanding of some items of the test, own statement that he had recently used drugs.

All of these symptoms, including a statement of the recent drug use, are only indirect evidence that the speaker is in a drug intoxication. Therefore, some record marked as belonging to a drug intoxicated speaker, may belong to a sober speaker, and vice versa. Nevertheless, the conclusions made on these samples are statistically valid.

The time of using drugs was estimated taking into account the following factors:

- Statement of the speaker that he is registered in the drug treatment clinic with a certain year
- Statement of the speaker that he is using drugs from a certain year
- Statement of the speaker that he was infected with hepatitis B or C in a certain year
- Statement of the speaker that he was infected with HIV in a certain year

Using all these factors we can only roughly estimate the duration of drug use. Many speakers repeatedly treated for drug addiction, and they said they had remission (a complete rejection of drug use), including long-term.

Almost all records of the speaker's in severe drug intoxication are very quiet, with an average amplitude of about 1-2 thousands of bins. However, a characteristic, such as energy, depends much more on the distance from the microphone to speaker's lips than on the state of the speaker. Therefore, below this characteristic is not considered.

Histograms of the fundamental frequency of the females are shown in Fig. 2. These histograms don't show any statistically significant anomalies. Median of the fundamental frequency distributions is in the range 190-220 Hz, which fits in with allowable limits for normal female voices. With some stretch only one histogram can be regarded as an anomalous, which has a two-mode distribution with the modes at the points of 120 Hz and 225 Hz. However, according to experts, a few percent of the female's voices have a distribution with two modes.

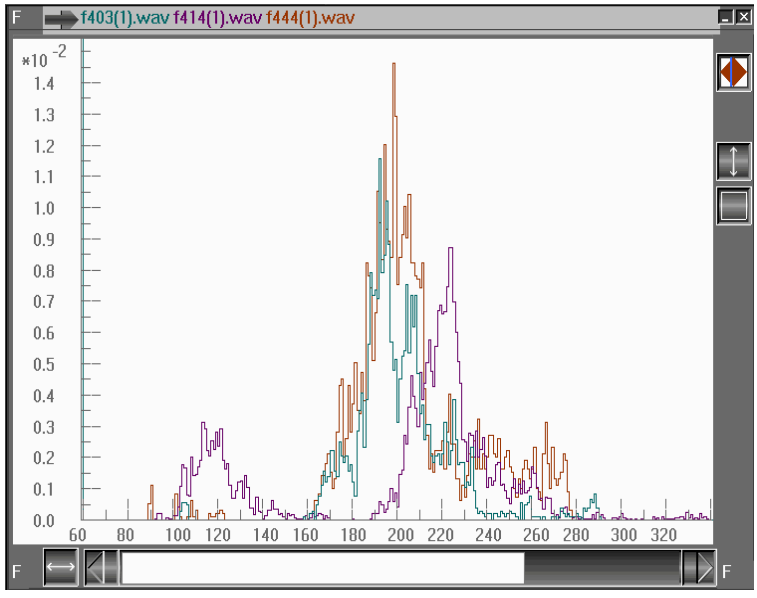


Fig. 2. Histograms of the fundamental frequency of the females in a state of severe drug intoxication

Fig. 3 shows examples of histograms on the fundamental frequency of 6 males in a state of severe drug intoxication. These histograms don't show any statistically significant anomalies. Median of the fundamental frequency distributions is in the range 110-130 Hz, which fits in with allowable limits for normal male voices. Only one histogram of six can be regarded as an anomalous that does not allow statistically significant conclusions.

A total out of three histograms of the female voices and 22 histograms of male voices in a state of severe drug intoxication were found only one abnormal histogram for female and 6 abnormal

histograms for males. Anomaly for men is in a low average fundamental frequency (below 100 Hz), which is rare in men not using drugs (according to experts - not more than 5% of men).

As a result, it may be concluded that the state of drug intoxication does not directly affect the fundamental frequency.

At the same time, the analysis of anomalies shows that of the 6 anomalies in men, in four cases, it takes place for the speakers who took drugs five or more years. The only anomaly in female voices belongs to the women who had used drugs within four years. This suggests another reason for the anomalies, namely, not the state of drug intoxication, but the length of time of using drugs.

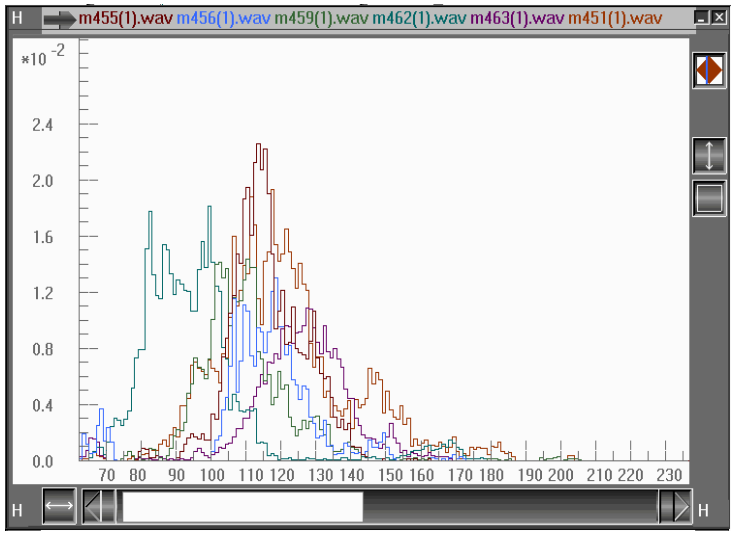


Fig. 3. Histograms of the fundamental frequency of six males in a state of severe drug intoxication

Fig. 4 shows histograms of the fundamental frequency of 15 males who use drugs five and more years. This is a completely different situation. The main mode for more than 2/3 of speakers moves from the 120-140 Hz range to 90-100 Hz range.

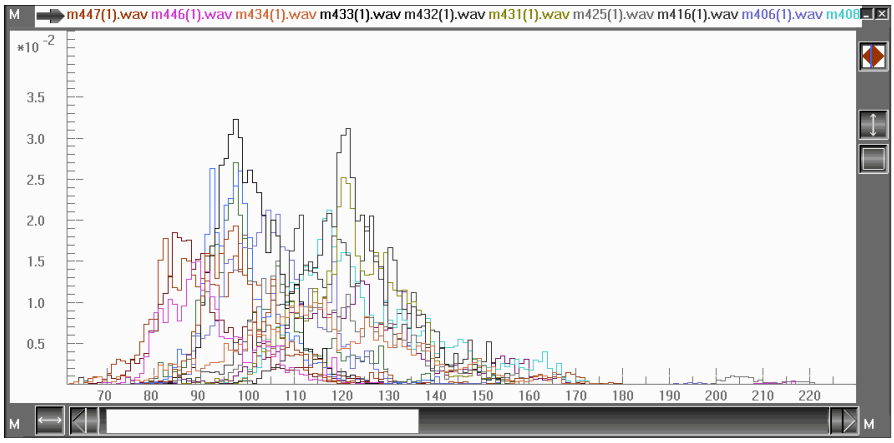


Fig. 4. Histograms of the fundamental frequency of 15 males who use drugs five and more years

The presence of the second mode at 120 Hz can be explained by the fact that the length of time of using drug is not known exactly, and obtained using indirect data. In addition, some speakers was treated for drug addiction, and they had a long-term remission.

If we sum all the histograms in Fig. 4, then after normalization we have the histogram in Fig. 5, where the mode of fundamental frequency distribution is at a frequency of about 97 Hz.

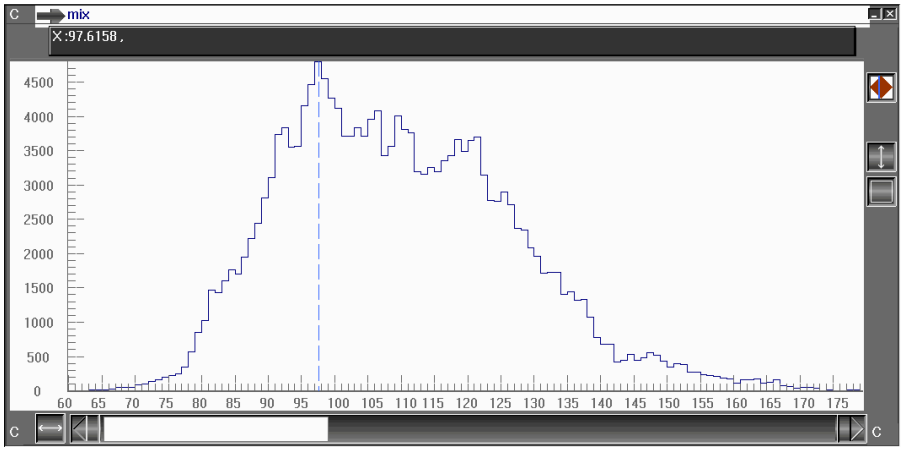


Fig. 5. Averaged histogram of the fundamental frequency of 15 males who use drugs five and more years

All this suggests that prolonged use of drugs of the heroin group leads to a decrease in the fundamental frequency.

The average rate of the decrease can be very roughly estimated as 4.5 Hz (3%) per year. The linear dependence of the estimate of the decrease is chosen as the most simple. Its applicability is limited to 2-8 years of intensive using of drugs of the heroin group. To study more complex dependencies and to evaluate the impact of a greater length of time using drugs obtained data are insufficient.

A similar conclusion for other types of drugs can not be done due to lack of data. The average rate of the decrease in the fundamental frequency may be somewhat underestimated and may be nonlinear. Analysis of types of drugs and the length of time of their use, suggests that drug users are gradually switching to heroin from "light" drugs.

Conclusion

Using the collected speech database with speaker's recordings in drug intoxicated and sober conditions and with different degrees of drug intoxication, we analyzed the speaker's voice before and after drug treatment and explored the effect of the duration of narcotic drug usage on the speaker's fundamental frequency.

The analysis showed that there are no statistically significant dependences in distortions of speaker's vocal parameters on different groups of narcotic drugs and the degree of drug intoxication.

Changes in the fundamental frequency are not regular and have no general nature. The main reason for changing the fundamental frequency is the speaker emotional state, rather than drug addiction treatment.

The study of the effect of the duration of narcotic drugs usage on speaker's voice showed that speakers with prolonged use of drugs of the heroin group have a tendency to decrease the fundamental frequency by about 3% per year.

References

1. *G.A. Adashinskaya, D.N. Chernov.* Acoustic correlates of person-distinctive functional and emotional states. Aviakosmicheskaja i ekologicheskaja meditsina, 41(2), 3–13, 2007

2. *H. Helfrich, R. Standke, K.R. Scherer*. Vocal indicators of psychoactive drug effects. Speech Communication, 3, 245-252, 1984

3. *V.Ya. Gelman, E.S. Dmitrieva, K.A. Zaitseva, A.M. Orlov*. Influence of Human Personal Features on Acoustic Correlates of Speech Emotional Intonation Characteristics. I.P. Pavlov Journal of Higher Nervous Activity, 59(5), 538-546, 2009

4. *F.J. Tolkmitt, K.R. Scherer*. Effect of Experimentally Induced Stress on Vocal Parameters. Journal of Experimental Psychology and Human Perceptual Performance, 12(3), 302–313, 1986

5. *P. Labutin, S. Koval, A. Raev*. Speaker identification based on the statistical analysis of f0. Proc. IAFPA 2007, The College of St Mark & St John, Plymouth, UK, 2007

6. *N. Smirnova, A. Starshinov, I. Oparin, T. Goloshchapova*. Speaker Identification Using Selective Comparison of Pitch Contour Parameters. Proc. ICPhS XVI, ID 1138, Saarbrücken, 2007

7. *Yu.N. Matveev, K.K. Simonchik*. The speaker identification system for the NIST SRE 2010. Proc. GraphiCon-2010, St. Petersburg, Russia, 315-319, 2010