Linguistics and Literature Studies 6(6): 278-284, 2018 http://www.hrpub.org
DOI: 10.13189/lls.2018.060603

Phonetic Cues Relevant to Drug Intoxication State Identification (Experimental Research)

Rodmonga Potapova[1,*], Tatiana Agibalova[2], Nikolay Bobrov[1], Natalya Zabello[1], Anastasia Yurashko[1], Ekaterina Migunova[1]

[1]Institute of Applied and Mathematical Linguistics, Moscow State Linguistic University, Moscow, Russia
[2]National Research Centre for Narcology, Moscow, Russia

Abstract This paper presents the results of analysis of perceptual auditory features enabling identification of drug intoxication states. The aim of the present investigation was to expand a set of personal characteristics of speech associated with the influence of drugs regarding Russian native female speakers. This paper presents preliminary results of the first stage of the investigation, the aim of which was perceptual auditory assessment of rambling speech (speech incoherency) by a set of parameters with regard to drug-intoxicated female speakers (opioid addicts) that can be regarded as distinctive features to be used for identification of drug-intoxicated speakers solely by their speech characteristics. At the present stage of the research, the perceptual auditory analysis method was used to single out the features relevant to drug intoxication identification. Further research is underway to establish automatically detectable acoustic correlates of these features.

Keywords Distinctive Phonetic Features, Drug Intoxication, Female Speakers, Forensic Phonetics, Perceptual Auditory Analysis, Segmental and Suprasegmental Features, Spoken Russian Language, State of Abstinence, Withdrawal Syndrome

1. Introduction

The subjective (additive) and objective (acoustic) events of speech behavior is of great interest for several professional groups. In the science of forensic phonetics the investigations in the field of acoustic and auditory features regarding voice and spoken language of subjects under the intoxication of drugs are very important. The aim of the present study is to add to the knowledge of several personal characteristics of speech under the influence of drugs regarding Russian native female speakers.

The term "personal characteristics of speech" [1] expresses the well-known fact that various speakers can be distinguished and recognized by their voices and speech. The personal characteristics of speech may be described as a complex of those sound qualities which enable us to identify the speaker. But our investigation is directed towards the problem of definition of those speech qualities, which are substantially a result of the psychic or psychosomatic conditions of a Russian female individual, of her emotional expression and of some somatic, neurological alteration in the region of her speech organs.

The study of disorders of spoken language is very important in the field of definition of the acoustic, physiological, psychic and psychosomatic characteristics of a subject. "It is important to identify the vocal correlates of these disorders (if they exist) for ordinary medical reasons... Another purpose for studying the voice/speech of psychotics is one that relates directly to the forensic milieu..." [2].

The theoretical frameworks [3] summarize the potential to guide the study of both normal and disordered speech (under the influence of drugs) [4, 5].

It should be noted that in our investigations we are using different scientific principles of the analysis in the field of forensic phonetics: auditory – acoustic – phonetic approach [4, 6, 7, 8, 9].

All speech disorders under the influence of drugs can be considered in relation to models of movement control (eg, monopitch / pitch level – bowed vocal folds, rigidity of laryngeal muscles; monoloudness / reduced loudness – rigidity of respiratory muscles) [3].

The complex auditory – acoustic – phonetic analysis of our Russian female speech corpus regarding speech signal in normal state and emotional state under the influence of drugs provides us a database of relevant individual characteristics of these speech signals. These results can be important for the determination of: a) relevant auditory and acoustic characteristics of Russian female speech (voice, prosody) under the influence of drugs; b) distinction of Russian female speech (voice, prosody) for normal state and without normal state.

The present research was conducted on the basis of a perceptual auditory experiment using the corpus of audio recordings of female native Russian speakers in the state of drug intoxication and abstinence. The hypothesis of this

investigation was as follows: all manifestations of psychosis or psychological stress can be interpreted as negative types of psychological states which are very similar to psychical mental diseases caused by drug intoxication states. From this point of view it is very important to investigate the drug intoxication states (this time of female Russian native speakers) by means of perceptual auditory and acoustic analysis on the basis of speech and voice characteristics as speech and voice correlates of psychotic manifestations. This assumption aligns well with a widely accepted hypothesis that drug intoxication, as well as other factors such as an emotional state and a functional stress, influence the person's speech and voice parameters [2, 4, 7, 10, 11, 12, 13, 14], but can also be illustrated by an opinion of Harry Hollien: "...the data and concepts about depression or affective disorders tend not to be as orderly as we would wish; indeed, they are rather difficult to organize. While the data suggest that depressive patients may utilize different speaking patterns than do normal, it also is possible that these patterns may not be universal or "classical" in nature" [2: 271]. The method of perceptual auditory analysis is based on the works of R.K. Potapova, the founder of the scientific school of fundamental and applied speechology in Russia [4; 5; 12; 15; 16; 17]. The investigation methods were elaborated in the course of ongoing cross-disciplinary research comprising a set of scientific problems, e.g. inter- and intraspeaker variability of prosodic parameters [4; 15; 16]; subjective and objective methods with regard to forensic speaker verification and identification [16; 17]; code switching identification and foreign language speaker recognition [16; 17]; emotional states identification [4]; verbal, paraverbal, nonverbal and extraverbal determinants of the personality [4]; cognitive ability testing in drug intoxication identification [5; 12]. In this research we investigated phonograms of speech produced in a state of drug intoxication and during drug-abuse treatment and show some perceptual features by which one can recognize whether the speaker is drug-intoxicated.

2. Materials and Methods

2.1. A corpus of phonograms (N = 130) of Russian speech produced by 40 female Russian native speakers in various drug intoxication states was recorded at the National Research Centre for Narcology, Moscow specifically for the research. All the subjects were aged from 20 to 45, having full school or higher education, of about the same social status (mainly office workers). The total duration of the phonograms amounts to 10 hours and 14 minutes. The phonograms contained various types of speech activity, namely, reading of specially selected phonetically difficult phrases, reading of a phonetically and cognitively rich text [15], retelling of a text, and spontaneous speech to allow for the most comprehensive evaluation of the differences in the drug-intoxicated speech from what is considered regular for Russian. In order to monitor the dynamics of their speech ability, most of the speakers were recorded three times:

- in the state of intoxication (usually, in the beginning of the hospital treatment);
- in the course of the treatment;
- in the end of the treatment (just before they left the hospital).

The time lag between the beginning, middle and end of the treatment varied from 15 to 25 days, i.e. there was usually a delay of 7 to 10 days between the recordings.
All the subjects were asked to fill in an informed consent form to satisfy all the applicable ethical and legal requirements.

2.2. A set of 23 phonograms (250 min 12 sec, 11 speakers in total) was selected for perceptual auditory experiments at the present stage of the research. The experiments involved listeners (N=24) who were native Russian speakers, Master's Degree students, both male and female, studying theoretical and applied linguistics. Each listener filled in a questionnaire reflecting his/her individual sociolinguistic background. The listeners were asked to listen to the recordings (as many times as necessary) and to write down the verbal content of the phonograms marking all phonetic disfluencies, irregularities and peculiarities (both segmental and suprasegmental) they perceived, including the range of pitch variation, melodic patterns, speech rate, pauses, voice strength, voice strength variation range, timbre, speech rhythm, differences in stressed and unstressed syllables in the phrase, speech breathing, and accuracy of pronunciation.

2.3. The list of irregularities that were expected to be found in the speech of drug intoxicated speakers initially included:

- filled (erm, eh... etc) and unfilled pauses;
- word/segment repetition;
- false starts;
- word interruptions;
- vowel lengthening;
- consonant lengthening;
- syllable lengthening;
- segment or syllable elision;
- word or word group ellipsis;
- segment, syllable or word insertion;
- syllable metathesis;
- substitution of unstressed (or weakly stressed) syllables for stressed syllables;
- equal stressing of all syllables;
- total absence of word stress;
- unnatural syllable accentuation;
- incorrect utterance segmentation into phrases and syntagmas;
- discrepancy between the communicative function of the utterance and its intonation (e.g., a complete

narrative phrase with interrogative intonation, or an imperative sentence with narrative intonation);
- variations in perceived speed of the speech (e.g. speech rate acceleration with regard to the whole utterance or some specific parts of it as opposed to steady speech rate);
- variations of perceived loudness of speech across various parts of the spoken text.

3. Results and Discussion

3.1. Preliminary analysis showed that the following phenomena are particularly frequent in the speech of drug-intoxicated speakers:
- hesitation pauses (both filled and unfilled);
- metathesis;
- replacement of consonant clusters with single consonants;
- consonant elision;
- false starts or syllable repetition.

On the whole, the segmental level appeared to be affected by the changes caused by intoxication to a greater extent than the suprasegmental level, and yet it is the prosodic changes that seemed to be most indicative of the intoxication state.

At the lexical level we could observe poor vocabulary, the use of fillers and introduction of words and syllables to phrases that did not contain them.

At the syntactic-semantic level: use of simple sentences, and theme-rheme contradiction.

At the logical-semantic level: lack of logical links, simplification of associations.

At the pragmatic level: loss of the communicative goal, cognitive impairment.

Diagram in fig. 1 shows the dynamics in the quantity of mistakes in phrases of a text being read as the average results for the drug-intoxicated speakers that were drug-addicts and those who had not developed addiction.

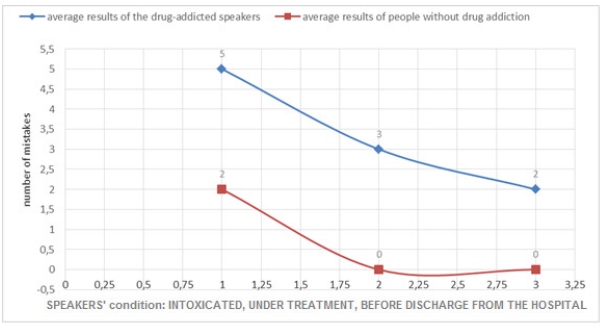

Figure 1. Dependence of the overall number of pronunciation mistakes in phrases of a text being read on the physiological condition of the speaker (drug intoxication, under treatment, after treatment) for speakers with and without addiction.

Among all mistakes in the text reading the following was singled out:
- Increased frequency and length of pauses, hesitation pauses.
- Indistinct speech rhythm, indistinct pronunciation
- Monotonous reading of the text, wrong phrases and segmentation.
- Unusual intonation and melodic pattern of the speech.
- Ellipsis, changes in the word syllable order, introduction of extra sounds, syllables and words (epenthesis).

3.2. The analysis of the phonograms for text retelling showed that the speakers used attributive and illustrative examples connected with their drug experience. Moreover, it was hard for the speakers to remember the content of the text; they tried to generalize it without giving details. For example, when retelling the text containing a description of a boat accident during a rowing competition they stated that “just something bad” happened during the competition. Another observation made was that the lexical complexity of the speech product of the drug-addicted speakers became uneven. The subjects tended to use simple sentence patterns, fillers, long hesitation pauses. As to the cognitive sphere, including semantic memory, the results showed the decrease (drop) in the level of the quality of generalization and abstraction and yet they sometimes retain certain words from the text belonging to a stylistic register that is high above the overall level of their speech (e.g. the word “испытания” ~ “ordeal”).

The results displayed in the second diagram (see fig. 2) show that the ability to render the text in a right way is restored during the drug treatment. In the first health condition the accuracy of the rendered text amounts to 35%, whereas the accuracy in the third condition increases up to 87%. The speakers used definite, shallow and primitive sentences, simple constructions, fillers and hesitation pauses. The subjects often omitted endings and left phrases and thoughts incomplete.

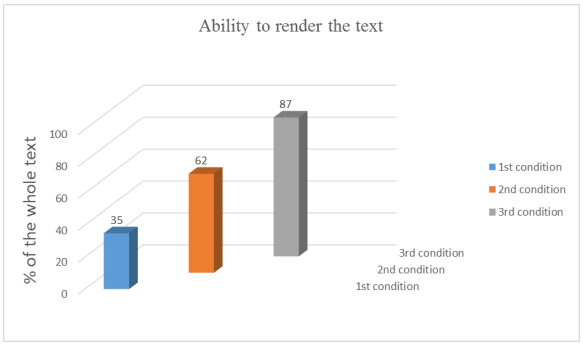

Figure 2. Dependence of the overall text rendering quality (% of adequately rendered sentences) on the physiological condition of the speaker (1 – intoxication, 2 – treatment, 3 – before discharge from the hospital).

3.3. Considering the data on the speech irregularities of drug-addicted speakers and the obtained statistics, one can

identify some general tendencies.

The majority of subjects have the voice pitch ranging from low to medium (fig. 3), which reflects their physical and mental state. Such tendency can also be partially explained by the influence of nicotine addiction that accompanied drug addiction in a vast majority of cases. The voice melodics can be characterized as monotonous, calm, "cheeky". In the 2nd and 3rd conditions (under and after treatment), the voice strength varied from weak to medium, but there was still a residual effect of drug intoxication, as well as the effect of medication. Comparison of the phrases and text pronounced in various conditions showed that the voice pitch variation remained within the same narrow range. The improvement of the physiological condition changed the melodic pattern from a monotonous one that had prevailed in 90% of the speakers, to a smooth one. The contour of the voice variation range in excitement was in the intermediate state.

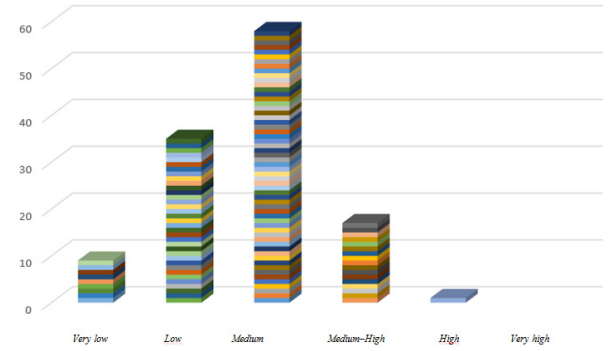

Figure 3. Perceived pitch distribution (in %).

The narrow pitch range prevailed, which accounts for emotional exhaustion as well (see fig. 4).

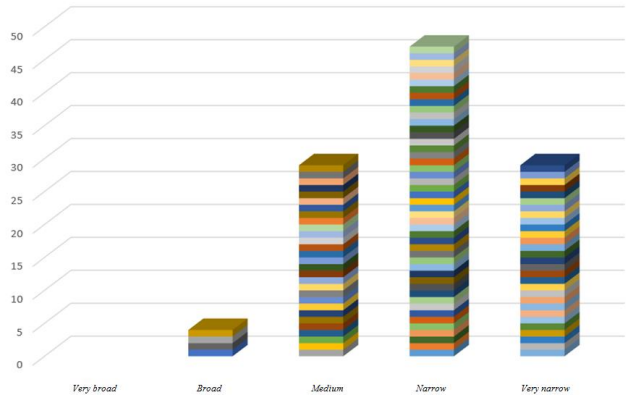

Figure 4. Perceived pitch range distribution (in %).

The dominating melodic pattern of the speech appeared to be abrupt and monotonous (see fig. 5). While reading the text and phrases the subjects, as a rule, did not pay attention to the degree of word emphasis and melodic pattern, due to which syllable chanting (equal emphasis on syllables) or lack of emphasis was observed.

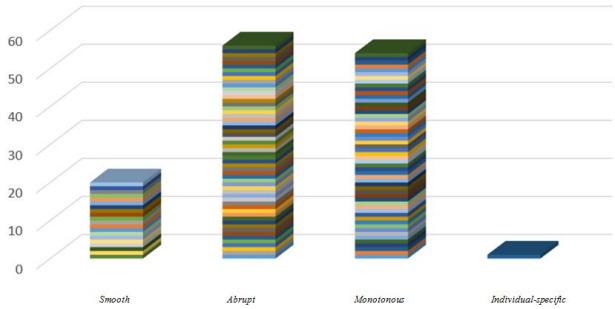

Figure 5. Melodic contour variation (in %).

3.4. The speech tempo is slowed down in most of the cases (fig. 6).

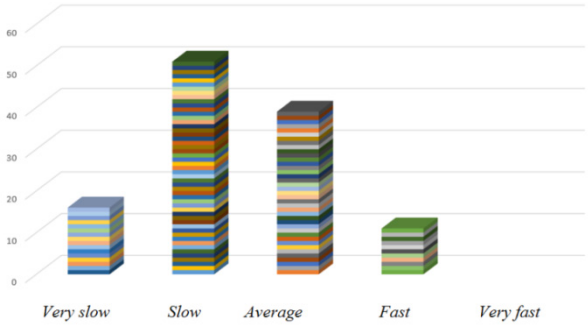

Figure 6. Perceived speech tempo distribution (in %).

The diagram of the frequency of pauses demonstrates that the listeners perceive pauses differently, hence it turned out to be quite difficult to estimate general tendencies here (fig. 7).

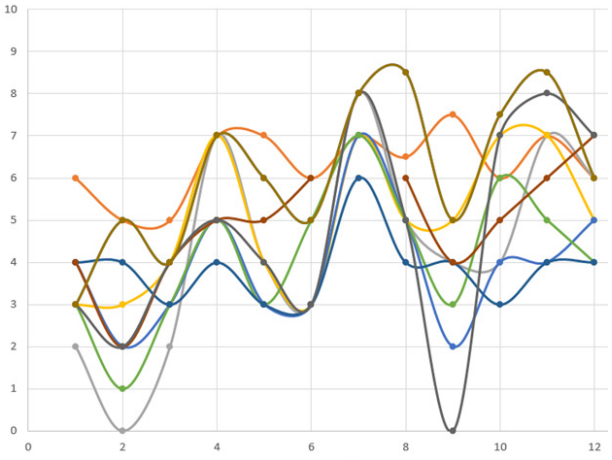

Figure 7. Frequency of pauses on the scale of 1 to 10 as assessed by various listeners (selected data)

Generally it can be noted that speakers made more pauses during the second and third interviews. It is worth mentioning that for every speaker without exception markers of control and correction were very typical. It indicates inhibited thinking, bradyphrasia, brain disorder and as a result various defects in perception, processing and reproduction of the information.

And yet it should be mentioned that not all the speakers follow the same pattern in what concerned speech tempo and pausing. Alongside the general tendency towards the slow speech rate, the elongation of the "gaps" between words, the tempo of some of the subjects was very high, while they made small number of very short pauses and nearly no hesitation pauses at all. The average duration of the pause interrupting the speech flow, i.e. the ratio of the total time of pauses to their number decreased in the subjects with the excitatory type of emotional tension, and increased in the subjects with the inhibitory type of emotional tension.

3.5. Most of the listeners indicated a hoarse, gruff and husky voice timbre evidencing an unpleasant experience during the listening (fig. 8).

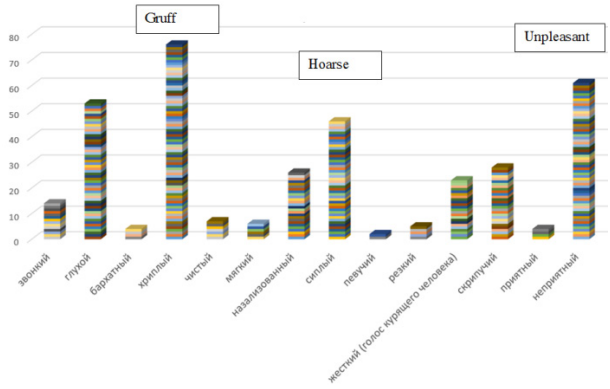

Figure 8. Voice timbre features perception (in %).

3.6. It is also quite obvious that irregular speech rhythm prevails (fig. 9).

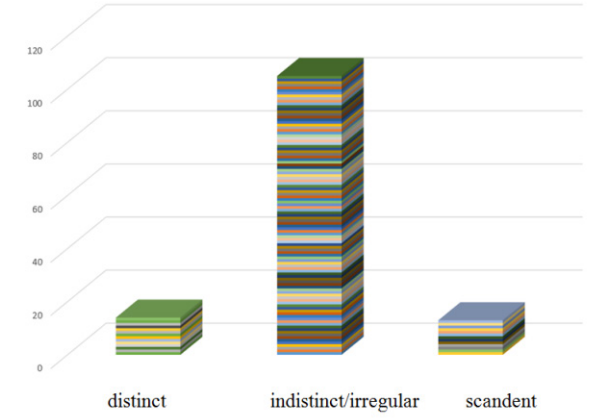

Figure 9. Perceived speech rhythm distribution (in %).

3.7. Another common characteristic in the majority of the speakers is irregular breathing or cogwheel respiration (fig. 10).

3.8. A prevailing characteristic is medium voice strength; however a large percentage of the speakers spoke in a very weak voice which also demonstrates a general tendency that once again points to physical and emotional exhaustion (fig. 11).

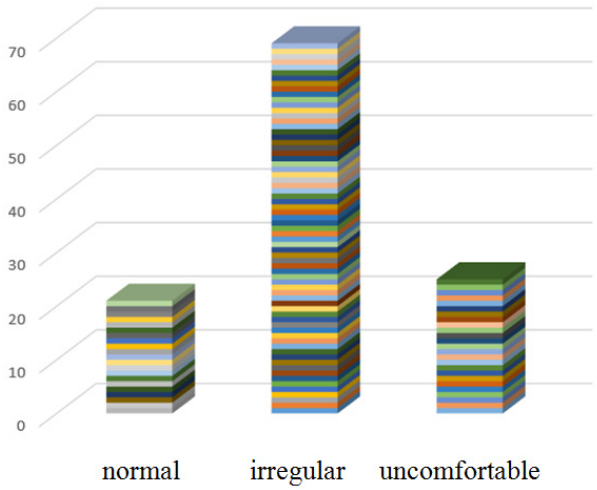

Figure 10. Speech breathing characteristics distribution (in %).

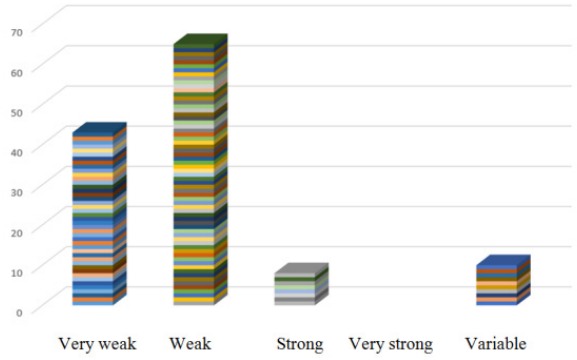

Figure 11. Perceived voice strength distribution (in %).

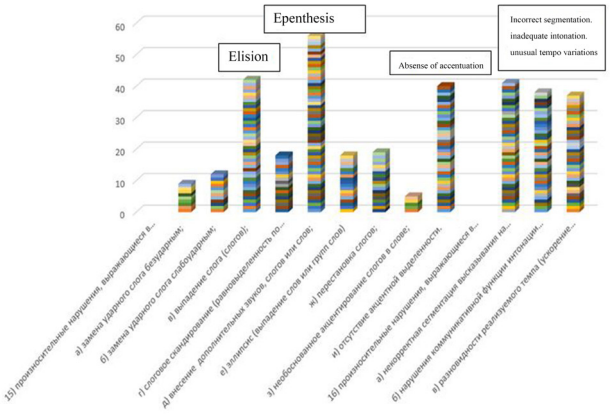

Figure 12. Relative frequencies of various speech irregularities in the speech of drug-intoxicated speakers (in %)

3.9. The diagram shows that introduction of some additional sounds into the speech is typical for all the speakers. Also some of the most frequent phenomena are dropped syllables, absence of accentuation in a sentence, incorrect segmentation of a sentence, irregularities connected with the communicative function of the intonation and variability of the tempo in a sentence (fig. 12).

4. Conclusions

Summing up the results obtained from the experiment the following conclusions can be made:

- Low and medium voice pitches along with the narrow pitch range prevail in the speech of the drug-intoxicated speakers.
- The melodic pattern of the speech is mainly either abrupt or monotonous.
- In most of the cases slow tempo prevails, however it can be pointed out that drug-intoxicated speakers sometimes had high speech tempo and short pauses, which can be explained by psychoemotional instability.
- All of the patients had hoarse, gruff and husky voice timbres. Besides this, the listeners pointed out to an unpleasant feeling while listening to the phonograms.
- Indistinct speech rhythm, "cogwheel" respiration and abnormal speech breathing are typical of the speech of the subjects.
- Medium voice strength is a prevailing characteristic; however a large percentage of the subjects spoke in a very weak voice.
- In spontaneous and quasi-spontaneous speech (retelling) the speakers made a lot of hesitation pauses. The second and third interviews with the patients demonstrated a great number of pauses, which can be attributed to gradual intoxication of the organism and long-term effects of the drug on the brain.
- Introduction of some additional sounds into the speech is typical for all the subjects. Also some of the most frequent phenomena are dropped syllables, absence of accentuation in a sentence, incorrect phrase segmentation, irregularities connected with the communicative function of the intonation and variability of the tempo.
- A great number of markers of control and correction appear in the speech of every speaker.
- All these general tendencies reveal a strong emotional and physical exhaustion, and also show the irreversible effect on the nervous system and the function of all the systems in general, primarily on the brain.

The results obtained are currently being used in the research aimed at the establishing of automatically detectable acoustic correlates of the mentioned features.

Future application of the results implies primarily the development of algorithms for automatic detection of drug intoxication indicators in the speech of Russian speakers, which is expected to be useful for emergency call services as well as risk assessment procedures in business decision making.

Acknowledgements

The research is supported by the International Association for Forensic Phonetics and Acoustics (head of the project – Sc.D., Prof. R.K. Potapova, full member of the IAFPA since 1997).

REFERENCES

[1] Janota, P. Personal characteristics of speech. Academia, Praha, 1967.

[2] Hollien, H. The Acoustics of Crime: the New Science of Forensic Phonetics. New York: Plenum, 1990.

[3] Kent, R.D. The acoustic and physiologic characteristics of neurologically impaired speech movements. In: Speech Production and Speech Modelling. Ed. by William J. Hardcastle and Alain Marchal. Series D: Behavioural and Social Sciences. Vol. 55. Kluver Academic Publishers. Dordrecht – Boston – London. Published in cooperation with NATO Scientific Affairs Division. 1990. P. 365–401.

[4] Potapova, R.K., Potapov, V.V. Language, Speech, Personality. – Moscow, 2006. (In Russian)

[5] Potapova, R.K., Potapov, V.V., Lebedeva, N.N., Agibalova, T.V. Interdisciplinarity in speech polyinformativity research / Potapova R.K. (Ed.). – Moscow, 2015. (In Russian)

[6] Kuenzel, H.J. Sprechererkennung. Heidelberg: Kriminalistik Verlag, 1987.

[7] French, P. An overview of forensic phonetics with particular reference to speaker identification. The International Journal Forensic Linguistics: Speech, Language and the Law. Vol. 1, issue 2, 1994. – P.169–182.

[8] Gold, E., French, P. International practices in forensic speaker comparison. The International journal of speech, language an the law. Vol. 18, No. 2, 2011: 293–307.

[9] Jessen, M. Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs. LINCOM Studies in Phonetics, 2012.

[10] Klasmeyer, G., Sendlmeyer, W.F. The Classification of Different Phonation Types in Emotional and Neutral Speech. Forensic Linguistics, 4(1), 1997. – P.104–124.

[11] Hollien, H. Forensic Voice Identification. New York: Academic Press, 2001.

[12] Potapova R., Potapov V. Semantic field of "drugs". – Moscow, 2004. (In Russian)

[13] Harrington, J. The Phonetic Analysis of Speech Corpora. Institute of Phonetics and Speech Processing, 2010. – URL: http://www.phonetik.uni-muenchen.de/~jmh/research/pasc010808/pasc.pdf

[14] Neuhauser, S. Phonetische und linguistische Prinzipien des forensischen Stimmenvergleichs. Review of the monograph by Michael Jessen (2012): LINCOM Studies in Phonetics (247 pp.). The International Journal: Speech, Language and the Law. Vol. 20.2, 2013. – P.325–330.

[15] Potapova, R.K. Speech: communication, information, cybernetics. 6 ed. – Moscow, 2016. (In Russian)

[16] Potapova, R., Potapov, V. On Individual Polyinformativity of Speech and Voice regarding Speakers Auditive Attribution (Forensic Phonetic Aspect). In: Ronzhin, A., Potapova, R., Nemeth, G. (Eds) Speech and Computer. Proceedings of the 18th International Conference SPECOM'2016, Budapest, Hungary. LNAI 9811. Springer International Publishing, 2016. – P. 507–514.

[17] Potapova R., Agibalova T., Bobrov N., Zabello N. Perceptual auditory speech features of the drug-intoxicated female speakers (preliminary results). Proceedings of the 26th Conference of International Association for Forensic Phonetics and Acoustics, Split, Croatia, 2017. –P. 118-121.