ACUSTEK
THE STANDARD IN PROFESSIONAL SURVEILLANCE SOLUTIONS

ACUSTEK Limited

Chamber Buildings, Suite 8,
54, North Street,
Swords, Co Dublin, Ireland

Business Centre "Morskoy vokzal"
6138, pl. Morskoy Slavy, 1
St. Petersburg, Russia

Expert's Report # 003-16

AUTHENTICITY EXAMINATION

of audio recording

"140124_0030-02.wav"

January 20th, 2016

St. Petersburg, Russia

About ACUSTEK

Located in Dublin, Ireland with additional office in St. Petersburg, Russia ACUSTEK is a supplier of professional products in the field of surveillance, technical protection and solutions for forensic audio investigation, providing a variety of forensic audio investigation services and training courses. Among our customers are the Ministry of Justice of Ireland, the Ministry of Defence and Internal Affairs of France, the Science Police and Carabinieri in Italy, Romania Special Services, Police Poland, Spain, the Security Service of Serbia, the Royal Gendarmerie of Morocco, Ministry of Defence and Internal Affairs of Egypt, China Police, Latvian Police, Expert Bureau of Georgia, solicitor's offices in the UK and Ireland as well as private individuals.

1. Expert

Ivan S. Siparov, dob 30/06/1979

Education:

St. Petersburg State University of Russia, Physics Department, Bachelor of Science (Physics, 2000), Master of Science (Physics, 2003) "Theoretical and mathematical physics".

Training Institution of Russian Ministry of Justice (Forensic Audio and Speech Examination, 2008).

Expert Competence Conformity Certificate of Russian Ministry of Justice (2008, 2013).

Work experience:

2006-2009: Speech Technology Center, Ltd. (Training manager, Forensic Audio Project manager, Expert), St. Petersburg, Russia

2009-present time: Forenex, Ltd (Senior Expert), St. Petersburg, Russia

2012-2015: Southern State University of Russia (Tutor and Lecturer for Master program "Theory and Practice of Forensic Audio Examination"), Rostov-on-Don, Russia

2012-present time: ACUSTEK, Ltd. (Expert, Chief forensic officer), St. Petersburg, Russia

More than 300 speaker voice identification cases, more than 250 authenticity analysis cases, 42 trainings for Forensic Audio Examination for law enforcement agencies all over the world. Testifying in court for both prosecution and defence. Forensic audio examination software applications developer.

2. Equipment

HW: Laptop HP EliteBook 8740w Intel Core i7 Q720 1.6 GHz, 8GB RAM, sound card Creative X-Fi Surround 5.1 Pro, headphones AKG K240 Monitor.

SW: Forensic audio system "ACU-Expert option LAB" (comprising OT-Expert 6.0 with additional module for audio authenticity analysis TD-Expert 6.0) approved by Russian Ministry of Justice for forensic audio investigation, Adobe Audition, RightMark Audio Analyser, Audacity. OS Windows 7 Ultimate SP1 (64-bit).

3. Task

Expert is acting upon request from Atanas Tchobanov (Bivol.bg) for authenticity analysis of audio recording from file "140124_0030-02.wav".

4. Audio recording

File name: “140124_0030-02.wav”

File format: WAV. Recording format: PCM, 44100 Hz, 16 bit, Stereo, Size: 132300118 Bytes; Duration: 00:12:30.00; Effective frequency range: 15450 Hz; Application listed in file header: Lavf56.37.100 (libsndfile-1.0.24), MD5: 080D7C560F33100952C2010E30EEA270

5. Methodology

The main approach to authenticity analysis is the verification of integrity of acoustic event registered. The order and essence of consequent evolving events must correspond to the logical structure (critical listening), semantic, lexical, prosodic, phonetic and psychological circumstances (linguistic analysis), technical features of physical environment of the recorded scene and technical characteristics of appearing of a trace of acoustic event – the signal itself (instrumental analysis). Typically, all approaches are complementary. Additional attention is paid onto the fragments, where there is a coincidence of artifacts, revealed by different methods.

According to criminological theory revealing of fact of editing is implemented throughout detection of manipulation traces. Therefore, the expert work is based on the statement, that in case no editing traces were revealed during the investigation, the audio is intact – it corresponds to the acoustic event that took place at the scene – and, thus, is an authentic evidence.

The expert comes from the statement that the copy of file, containing digital audio recording does not introduce any distortions to

the audio signal, thus the file copy can be used to establish authenticity of recorded audio, that was originally created at the scene.

The following methodological background was used during the investigation:

Auditory analysis (critical listening):

- Analysis of continuity and simultaneity of audio acts and events, background acoustics;
- Search for any kind of artefacts that can appear after modifying of original audio data;
- Search for unjustified changes of number of speakers;
- Search for sudden changes of speakers' voice strength and timbre.

Linguistic analysis (in particular communication behaviour and phonetic analysis):

- Search for breaches of phonetic issuance of speech (syllables and words);
- Search for breaches of prosodic issuance of speech (phrases);
- Search for sudden changes of speakers' speech behaviour.

Instrumental analysis:

- Inconsistency of audio data to the technical parameters of the recording device and recording carrier (if provided);
- Search for the traces of anti-aliasing filtering;
- Search for the traces of analogue/digital start/stop/pause.
- Search for unjustified sudden changes of spectral and cepstral (autocorrelation) structure of the audio signal;
- Search for unjustified sudden changes of energy dynamics of speech signals;

- Inconsistency of reverberation time to the circumstances of the recording;
- Search for unjustified sudden changes of acoustic environment;
- Search for unjustified sudden changes of background noises in the time and spectral domain;
- Search for unjustified sudden changes of value, magnitude and phase of the technical signals (narrow-band parasite signals);
- Analysis of the continuity of speech signals or any periodic events;
- Search for unjustified sudden changes of the samples distribution;
- Search for unjustified sudden changes of the DC Offset behaviour;
- Search for Mp3, WMA, AAC, OGG codec traces by signal features and MDCT-spectrum;
- Search for codec frame borders disorder.

General conclusion is given on the summary of the auditory, linguistic and instrumental analysis considering the significance of each method in the particular circumstances.

Literature:

Contemporary methods, technical and software tools used for forensic audio examination. A.Sh.Kaganov. Forensic Examination Centre, Ministry of Justice, Russian Federation, 2003.

Search for the traces of editing and authenticity analysis of speech recordings. S.L.Koval. Speech Technology Centre. 2006.

Forensic authenticity analysis of analog and digital speech recordings. G.N.Zubov, P.I.Zubova, I.S.Siparov. Forenex. 2011

The Acoustics of Crime: The New Science of Forensic Phonetics (Applied Psycholinguistics and Communication Disorders, Harry Hollien, 1990

6. Investigation

The recording duration is 12 minutes 30 seconds. Signal waveform is represented at Fig.1

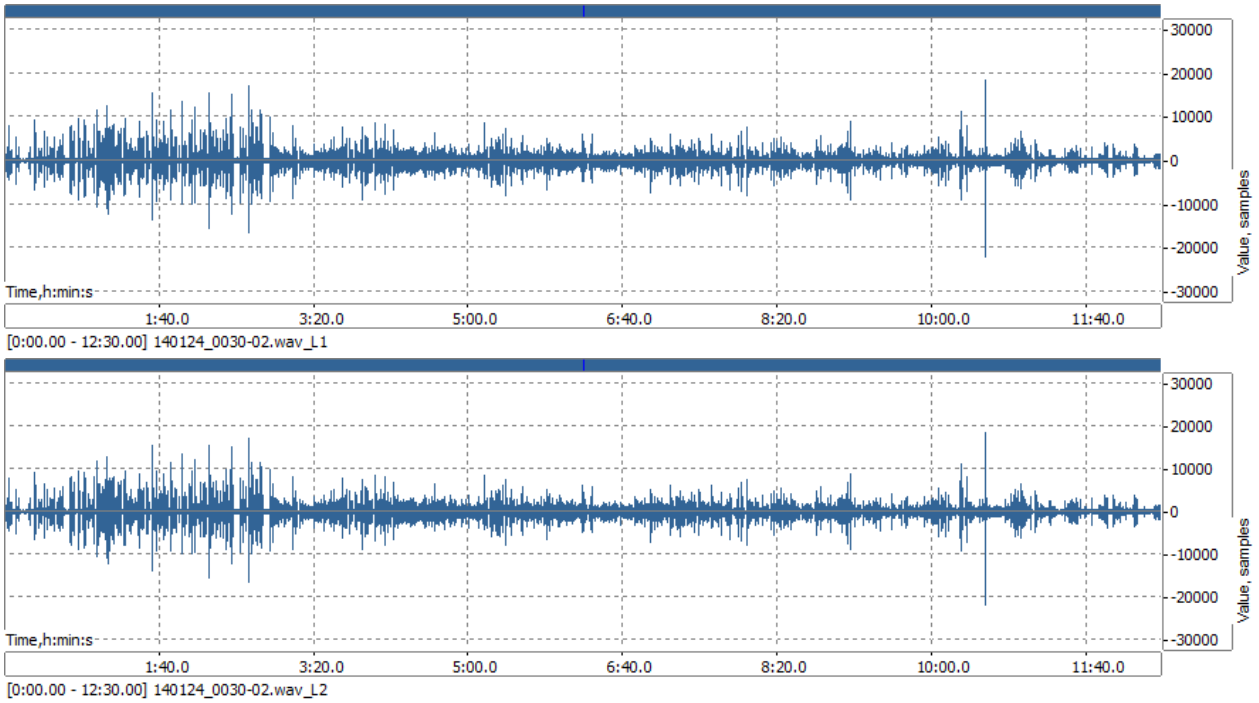


Fig.1. Waveform of left (top) and right (bottom) channels of investigated signal.

Analysis of continuity and simultaneity of audio acts and events, background acoustics did not reveal any editing or manipulation features. The recording background does not change during the time. The acoustic scene is quiet. There are sounds of crockery and cutlery. Also, there are sounds similar to a small dog (or some other pet) snuffing and barking very quietly.

Two speakers with female voices are participating in the conversation. The distance from speakers to microphone is different and does not change during the time. One speaker (louder, closer to mic) is giving a wide description, the other is listening and supporting the communicative act. There are no sudden and unjustified change of number of speakers, of their voices strength and timbre. Speech volume and audio recording quality allows to identify speakers.

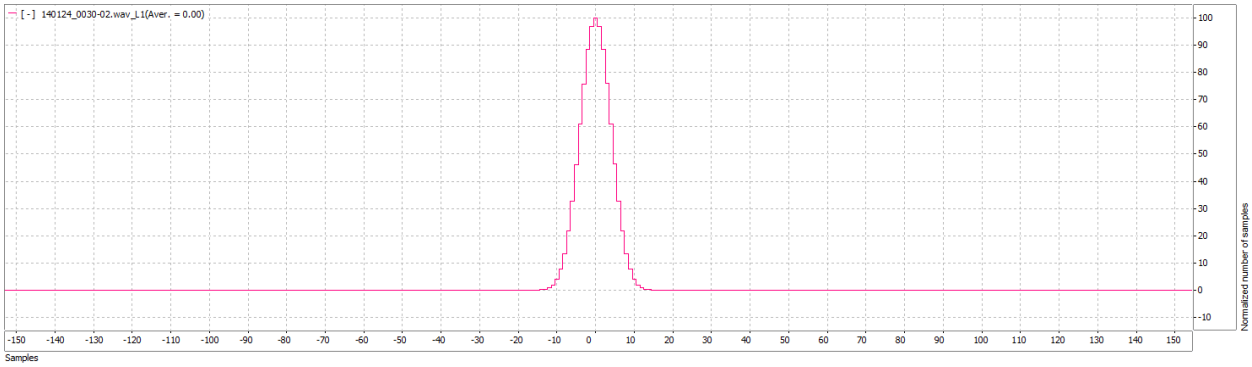
Phonetic and prosodic flow of the speech has no breaks. Speakers participate in communicative act as their communicative roles are equal – interrupting and replying. Phonetic speaker's profiles are not changing during time.

Due to the reason that spoken language is not native to or known by expert logical, semantic and thematical conversation unity could not be established. Lexical speakers profile permanence could not be established, too.

Speaker's speech behaviour is not changing during the recording.

Technical investigation established the following:

Signals from two channels were recorded by one or two very close microphones. The channel difference is less than ±15 samples (Fig.2). Samples histogram does not have traces of compression or manipulations (Fig.3).



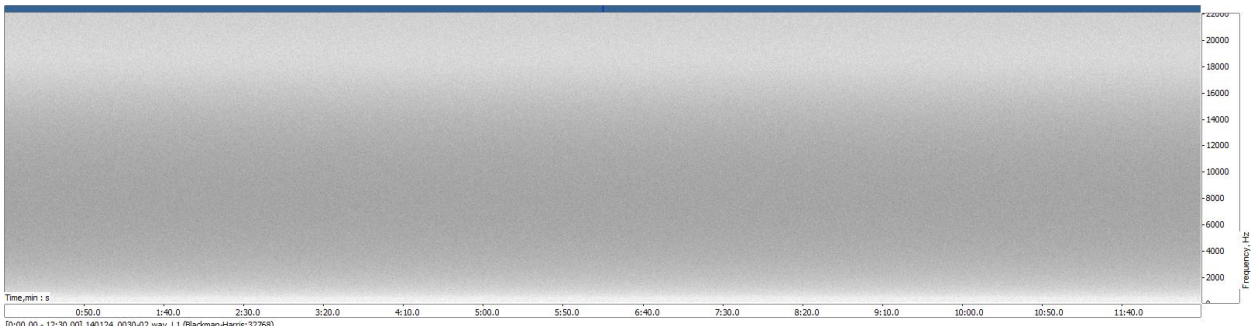


Fig.2. Samples histogram (top) and dynamic spectrogram (bottom) of channel signal difference. Channel difference does not change during time.

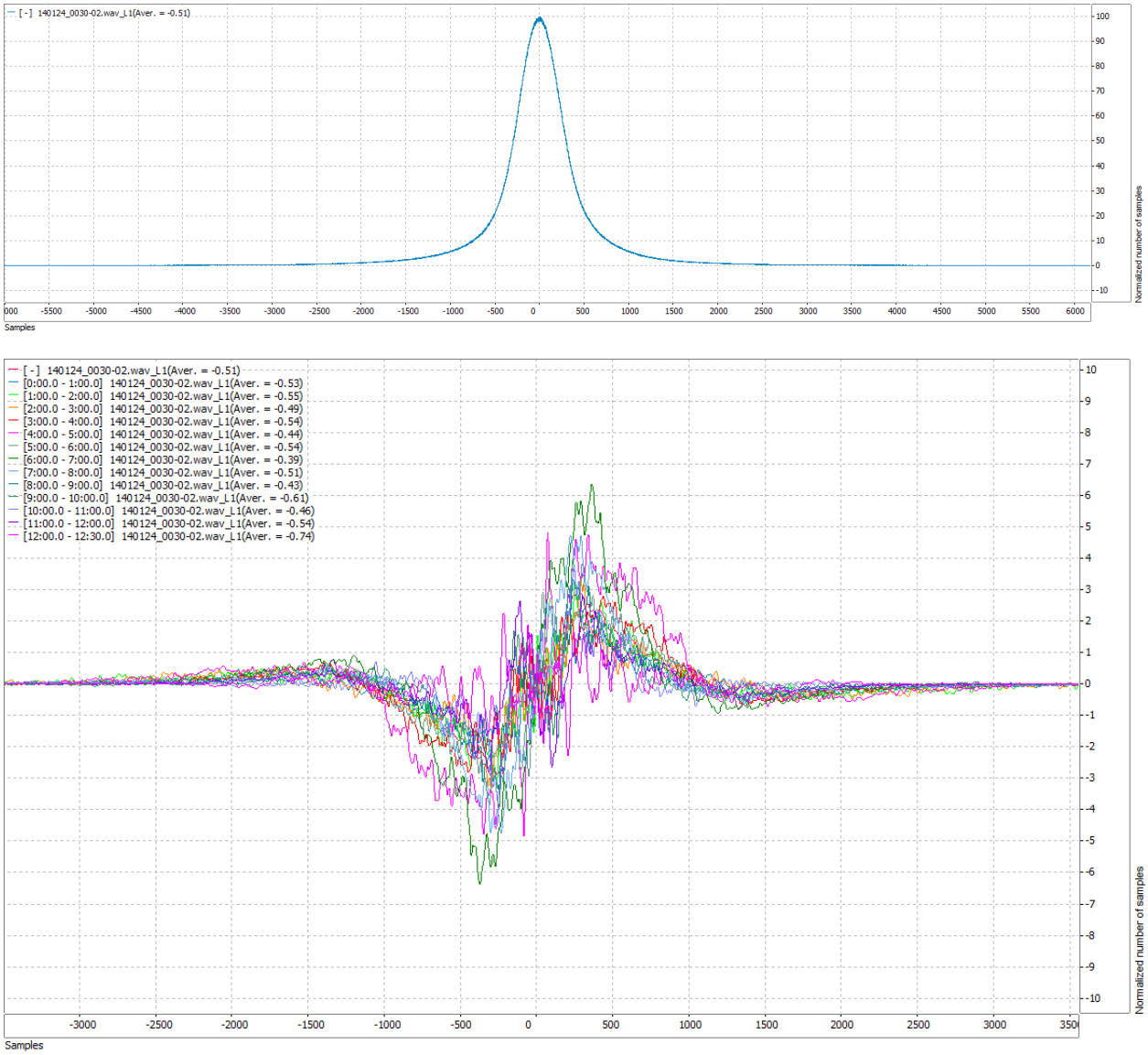


Fig.3. Samples histogram of the whole signal (top) and samples histogram asymmethry of consequent signal fragments (bottom, matches).

Signal energy changes smoothly, without sudden and sharp breaks (Fig.4). Every acoustic impulse has the non-zero starting front duration.

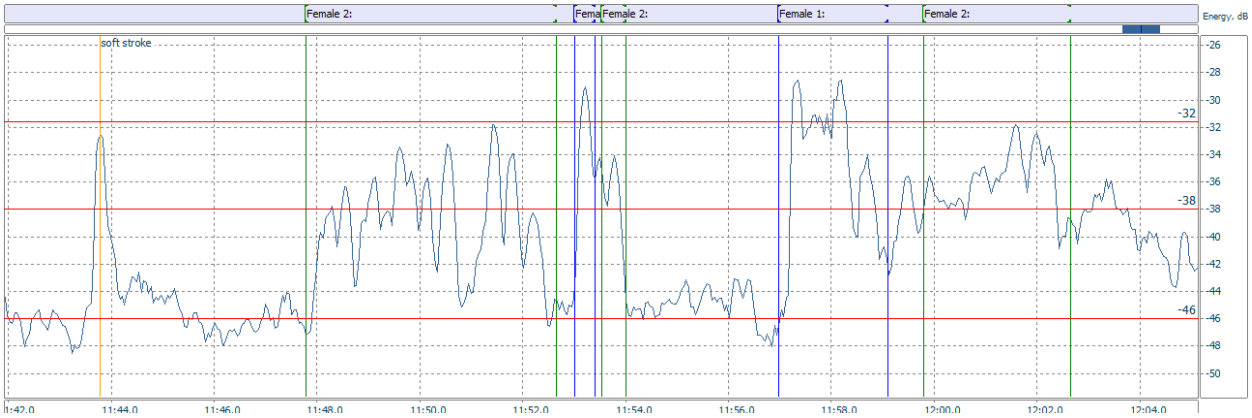


Fig.4. Signal energy. Green marks correspond to phrases said by "far" speaker, blue – for "close" speaker, yellow cursor mark soft stroke. Horizontal marks correspond to (down up) energy level of broadband noises, "far" speaker phrases, "close" speaker phrases.

Pulses, clicks and individual pronunciation features do not break the signal continuity (controlled by dynamic spectrograms and cepstrograms) (Fig.5).

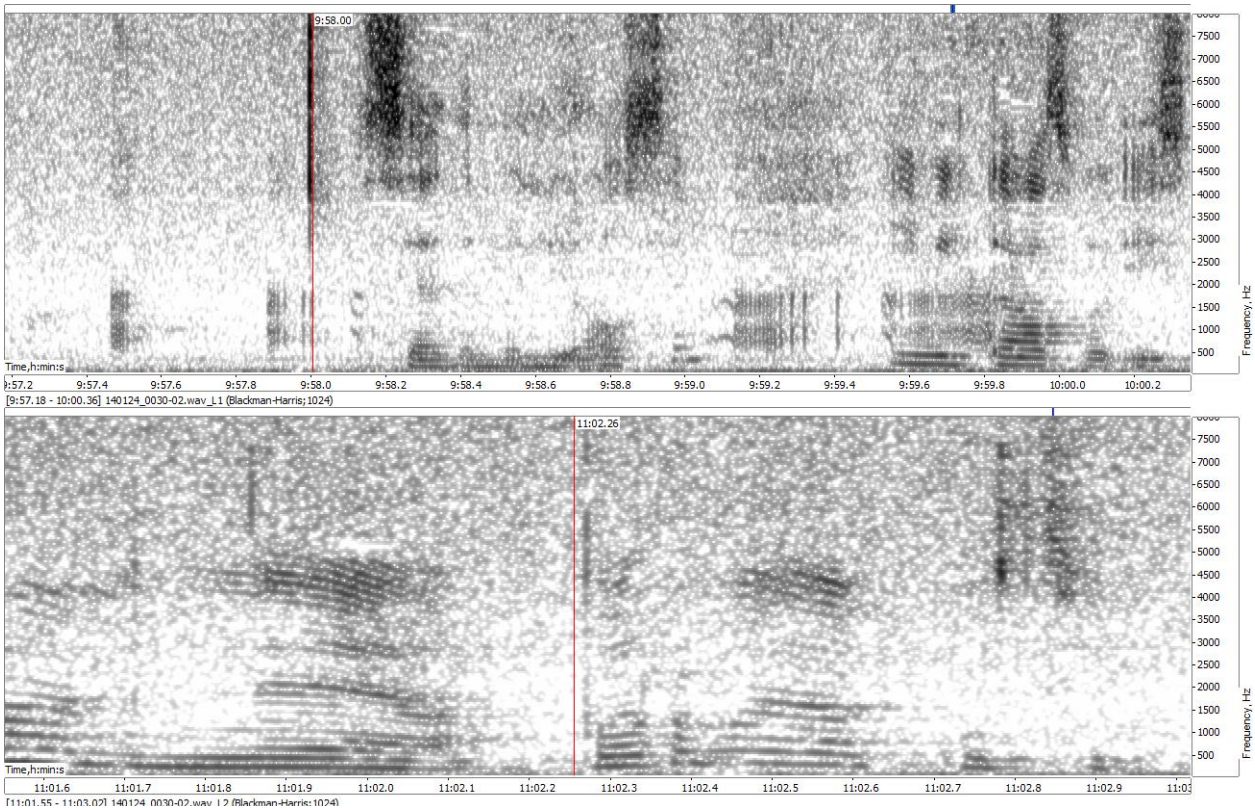


Fig.5. The click marked with cursor (top) does not change the spectrum shape. Realization of consonant [p] is depicted on bottom picture.

DC-offset of the signal does not contain traces of manipulations (Fig.6). There are no signal fragments with repeating samples (more than 4) or any traces of recording mode changing.

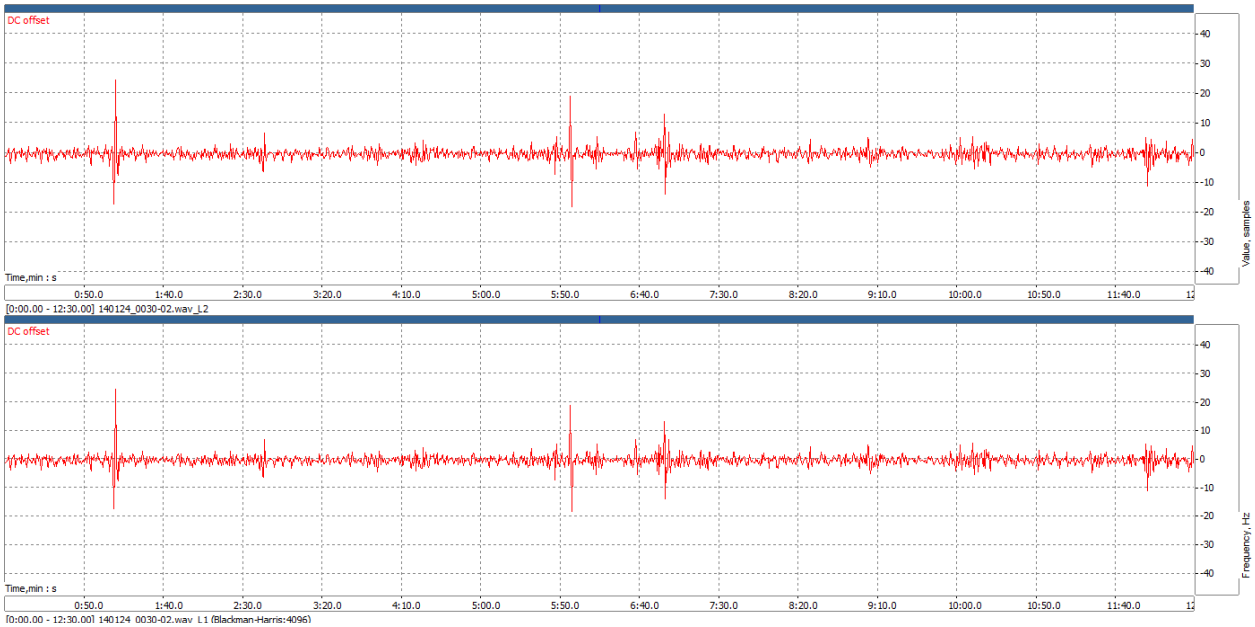
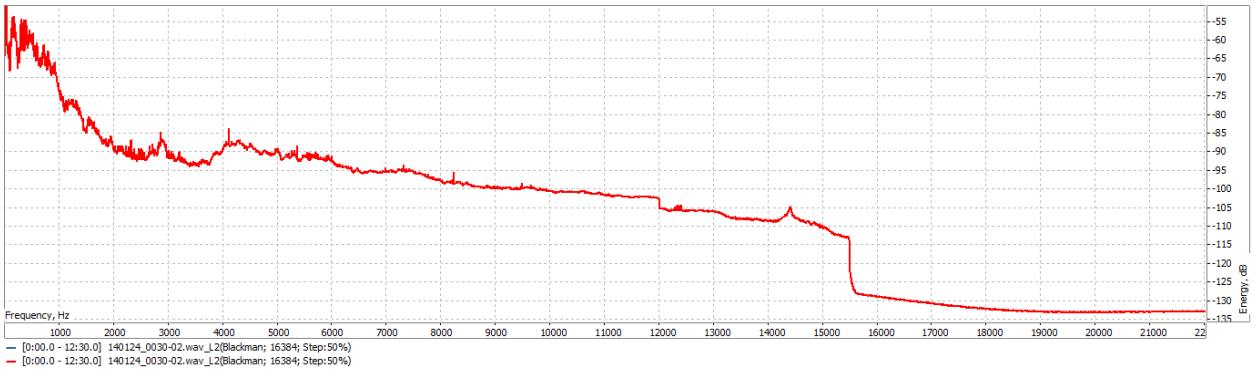


Fig.6. DC-offset of left (top) and right (bottom) signal channels.

There are no traces of anti-aliasing filter. Spectrum shape at 12000 Hz and 15450 Hz (Fig.7) is determined by audio recording format (psychoacoustic codec).



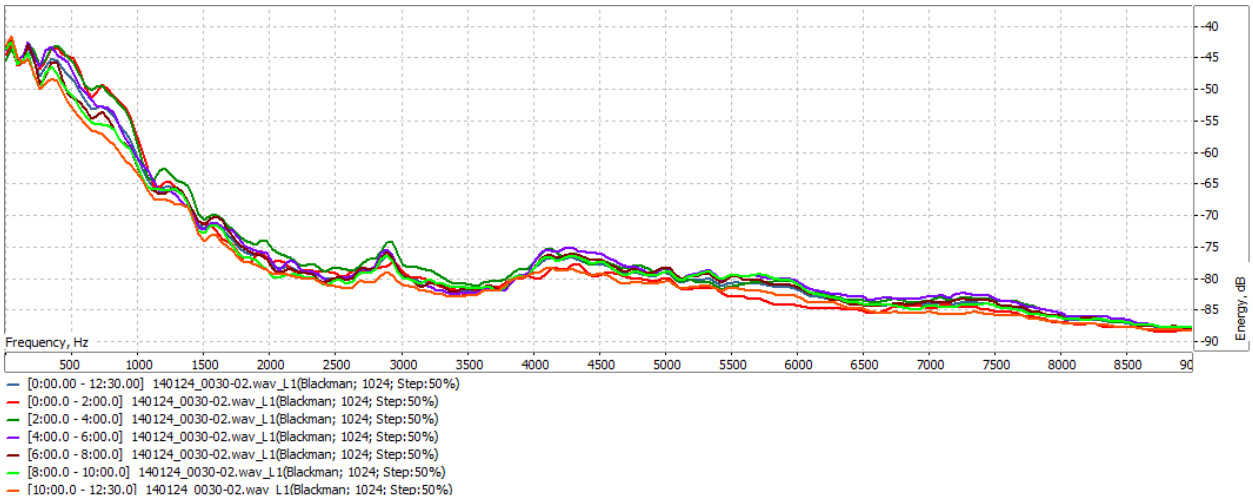


Fig.7. Average spectrum (top) of the whole signal (left and right channels spectra matches) and average spectra of signal fragments (do not cnange the shape during time) (bottom).

There are traces of psychoacoustic codec that are finely revealed over signal's dynamic spectrogram (Fig.8).

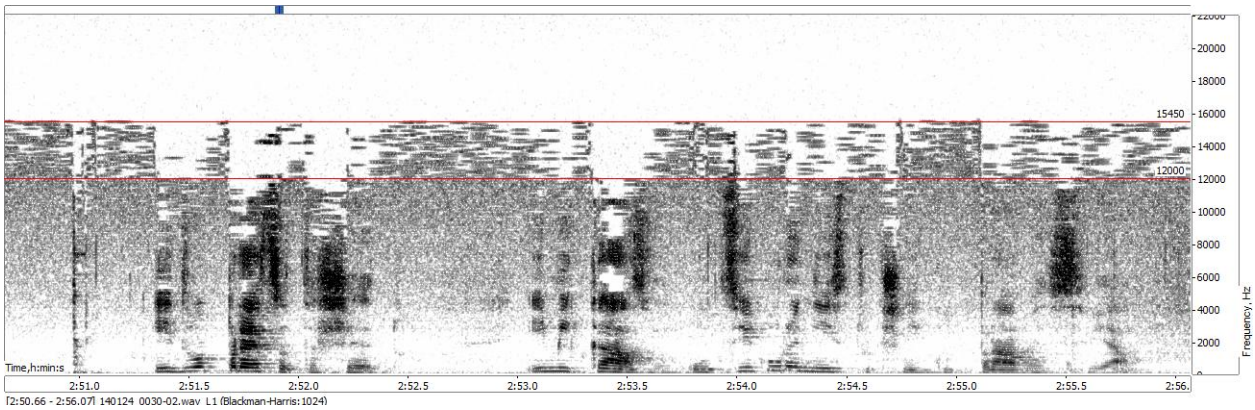


Fig.8. Dynamic spectrogram of the signal reveals traces of codec used to record the audio (rectangular areas of zero spectral amplitude).

MDCT based codec detector established that the signal was recorded with WMA or OGG recording format (Fig.9). Frame shift graph does not have breaks on the whole signal. Thus, the signal continuity is left intact. Example of MDCT-spectrum and histogram of frame borders positions is given at Fig.10.

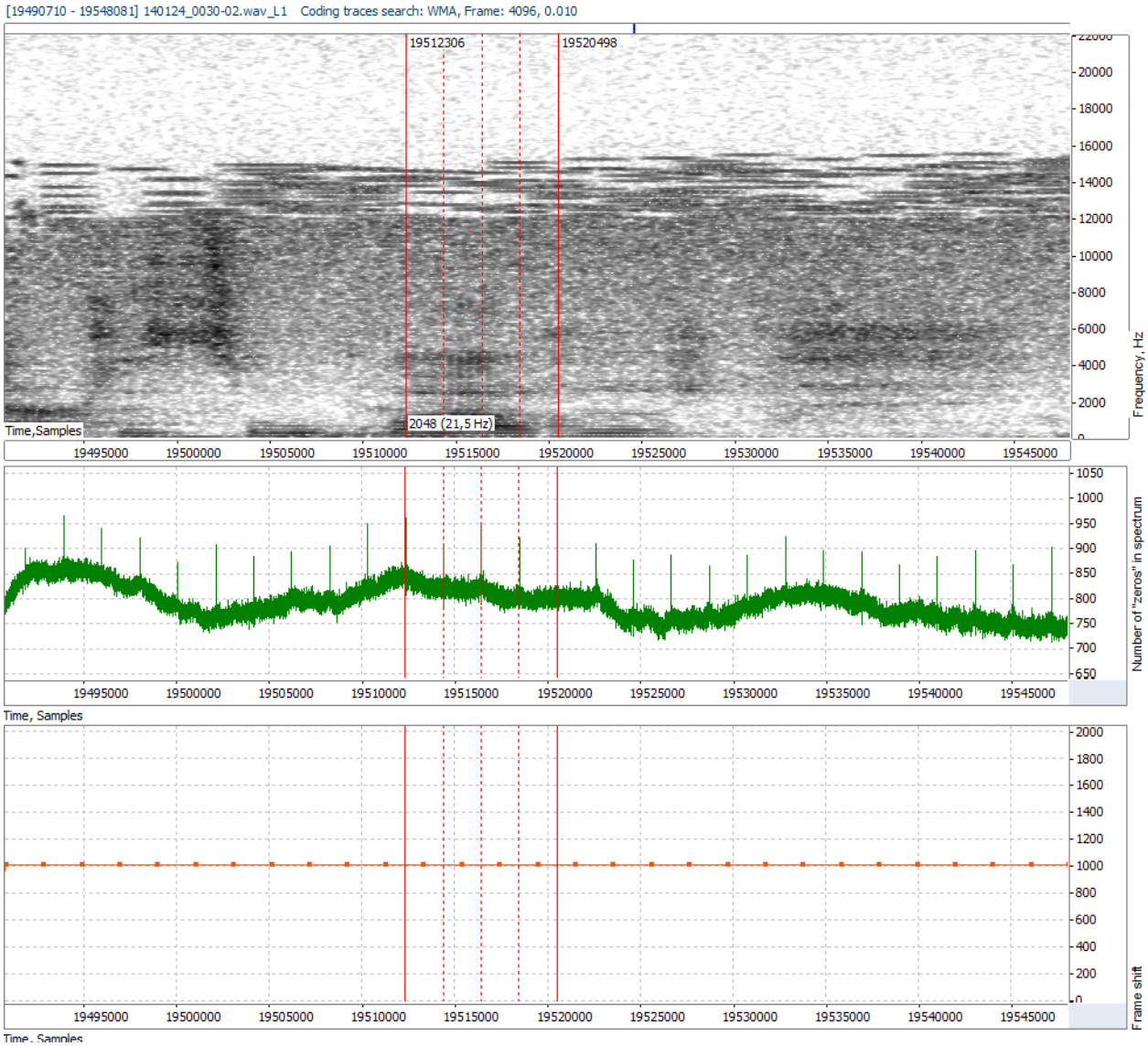


Fig.9. Dynamic spectrogram (top), number of MDCT-spectrum zeros (middle) and frame shift position (bottom). Vertical marks corrsepond to position of codec frame centers.

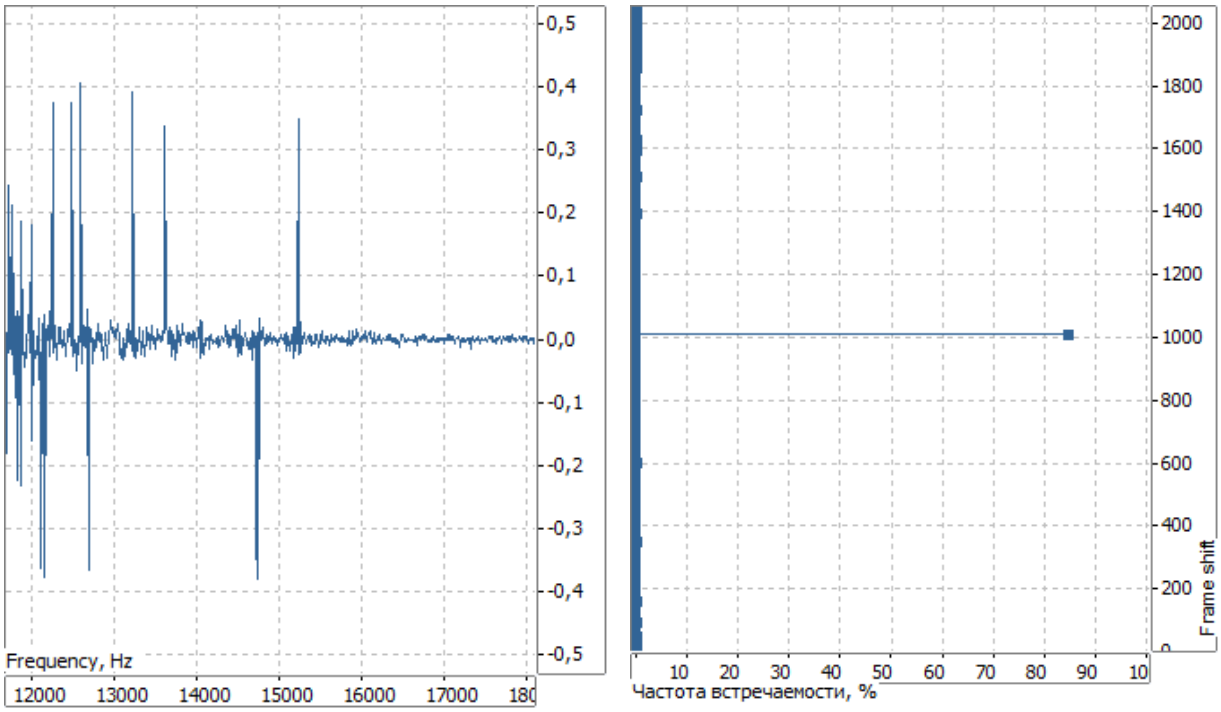


Fig.10. MDCT-spectrum (left) and frame borders histogram for WMA analysis (right) that approves signal was coded by psychoacoustic codec.

There are several narrowband technical harmonics in the signal that could be classified for two groups: stable and changing frequency during time. These harmonics can appear due to the parasite pickups of EM equipment or networks that belong to the recording scene circumstances.

There are three ENF harmonics in the signal (Fig.11), which correspond to the situation of signal coping through analog (acoustic or electrical) channel. Expert believe that coping itself cannot be regarded as a stand-alone feature of signal manipulating or its approval.

Particular origin of other harmonics cannot be established.

Phase continuity of harmonic 90.58 Hz is represented at Fig.12. Low relative level of other harmonics does not allow to carry out informative analysis of their phase. The moments of time, were harmonics 99 Hz and 100 Hz change their energy (Fig.13) with a surge were checked accurately with other authenticity analysis methods. The check did not reveal any other traces of signal continuity breaks.

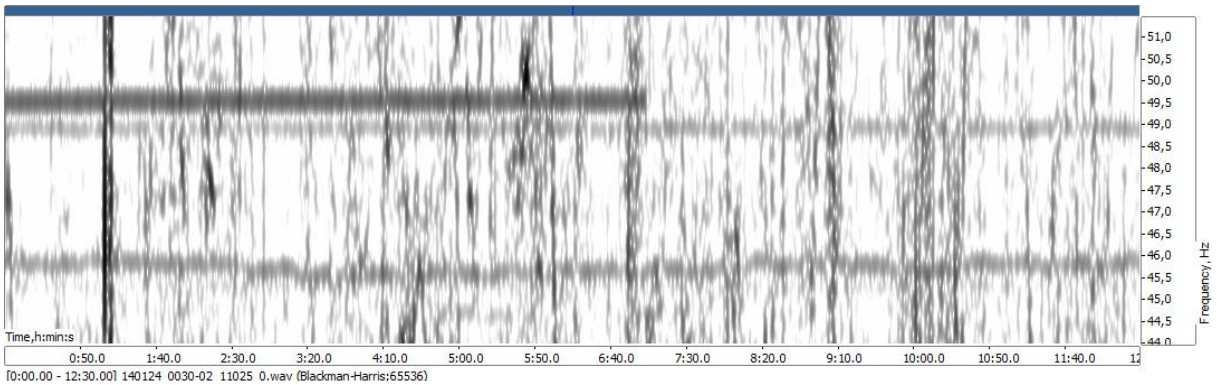


Fig.11. Dynamic spectrogram of ENF (50 Hz).

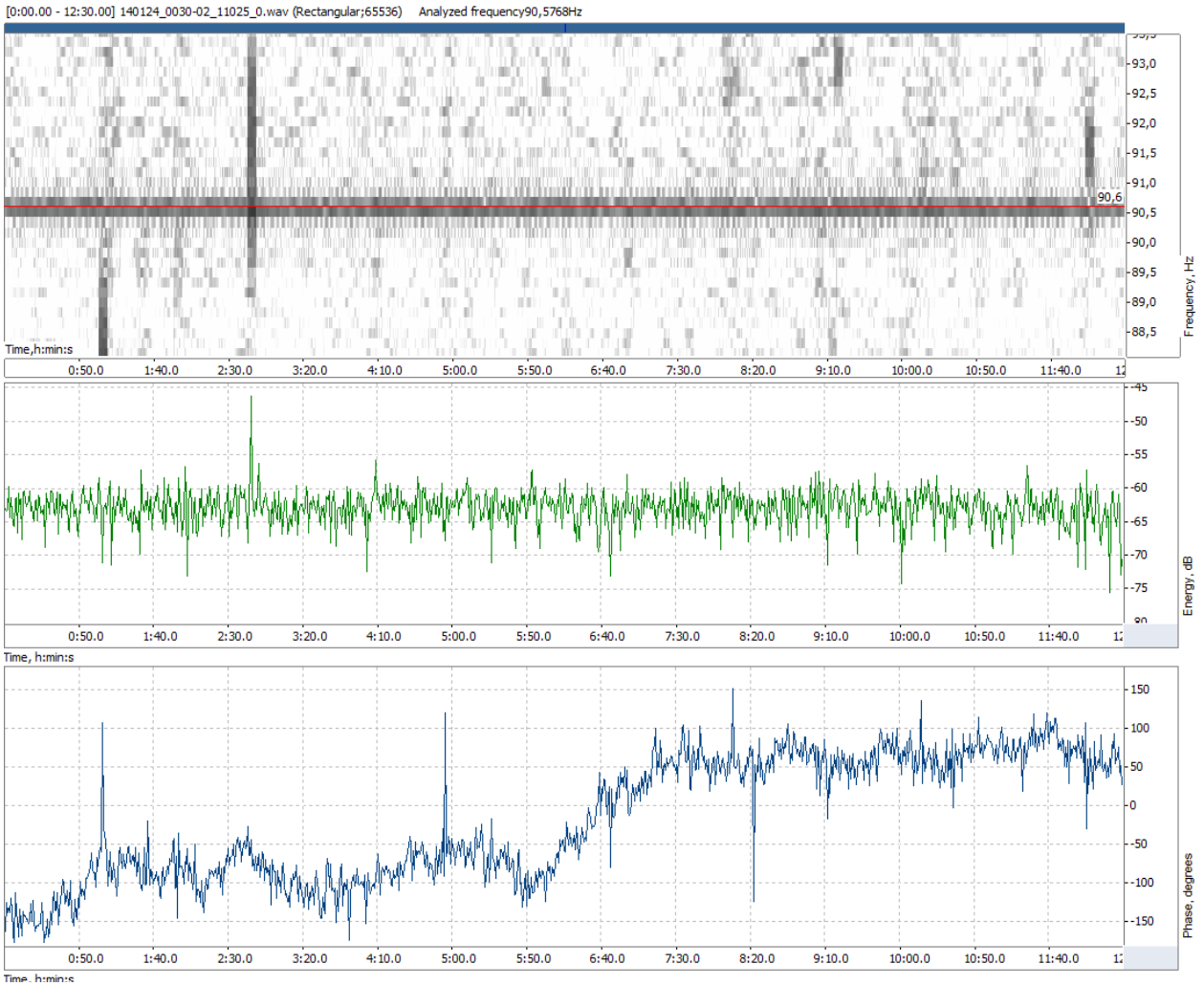


Fig.12. Dynamic spectrogram (top), harmonic's energy (middle) and phase (bottom). There are no traces that could approve signal editing.

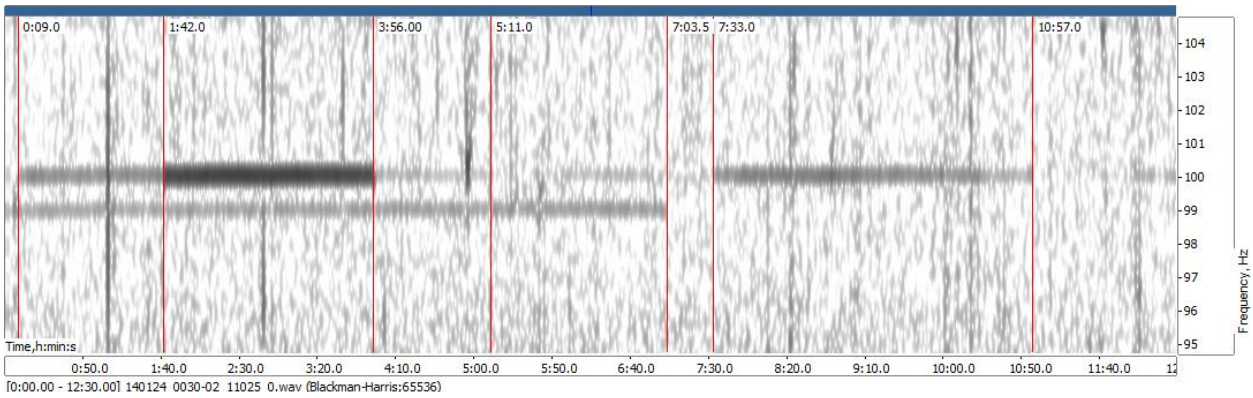


Fig.13. Dynamic spectrogram of 100 Hz frequency range. Vertical cursors correspond to sudden changes of harmonic's energy.

Narrowband harmonic with exponentially decreased value are represented at Fig.14. Frequency value changes smoothly, no informative breaks detected.

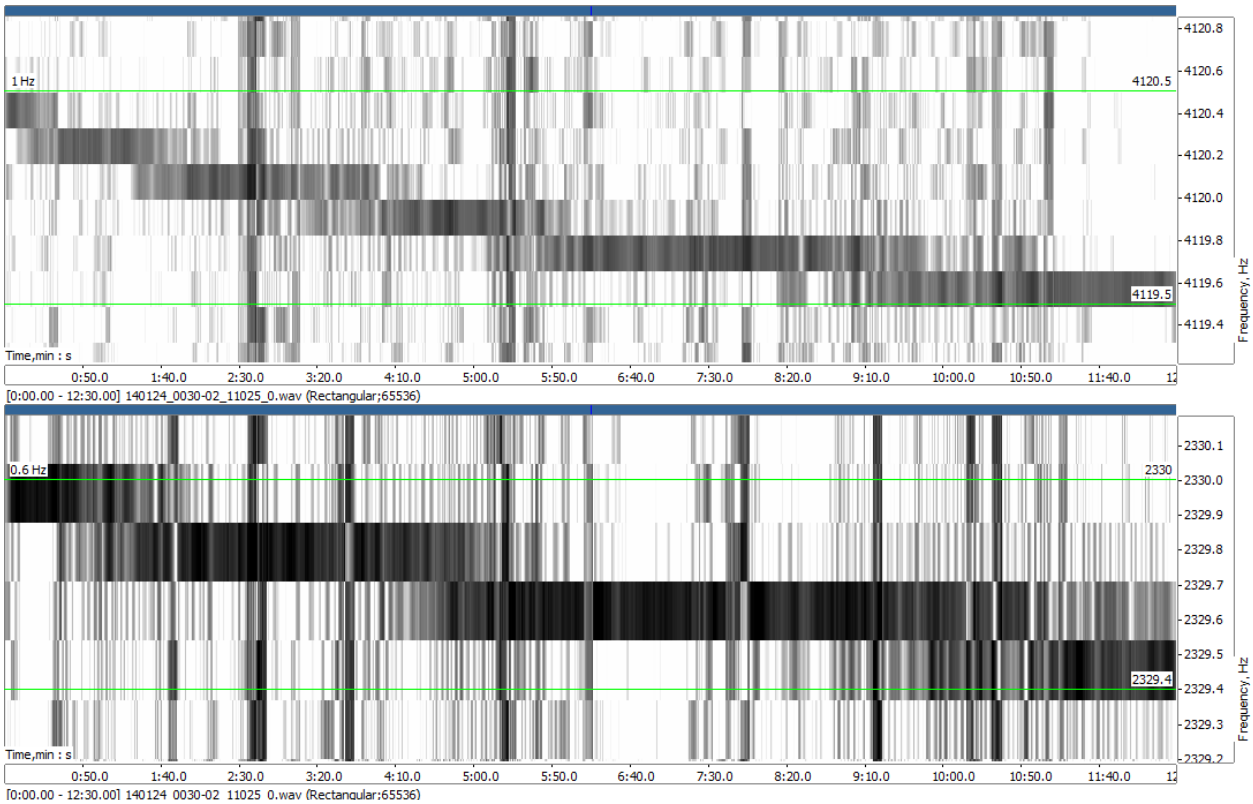


Fig.14. Examples of narrowband harmonics that are changing frequency value during time inside the range marked with green cursors.

Also, instrumental analysis did not reveal:

- unjustified sudden changes of spectral and cepstral (autocorrelation) structure of the audio signal;

- reverberation time to the circumstances of the recording;
- unjustified sudden changes of background noises in the time and spectral domain.

Therefore, summarizing the facts established during auditory, linguistic and instrumental analysis of the audio submitted for the investigation the expert comes to conclusion, that:

- the audio provided is the copy of the audio recorded at the scene, as soon as the initial audio contain traces of WMA or OGG codec used to record it;
- the audio provided has no breaks of continuity or traces that approve signal editing or manipulating;
- order and essence of acoustic events represented at the audio corresponds to those taken place at the scene.

7. Conclusion

Summarizing the facts established during auditory, linguistic and instrumental analysis of the audio submitted for the investigation the expert comes to conclusion, that:

- **the audio provided is the copy of the audio recorded at the scene, as soon as the initial audio contain traces of WMA or OGG codec used to record it;**
- **the audio provided has no breaks of continuity or traces that approve signal editing or manipulating;**
- **order and essence of acoustic events represented at the audio corresponds to those taken place at the scene.**

Expert: Ivan S. Siparov

Authorised by: ACUSTEK, Ltd.