

# Forensic Voice Analysis in Criminal Investigations<sup>1</sup>

*Jadranka Otašević<sup>2</sup>, Božidar Otašević<sup>3</sup>, Saša Atanasov<sup>4</sup>*

Forensics, as a multidisciplinary activity, has as its main task the provision of material evidence of all kinds. The scope of forensics includes all preliminary activities that are a necessary precondition for providing evidence, as well as work on clarifying criminal acts, especially those with unknown perpetrators, as well as actions performed immediately after cognition about the existence of the criminal act and which are part of police action. One of the main goals of forensics is the identification of persons. In recent years, there has been a growing interest and need to identify the perpetrator of a crime based on voice. A person's voice is a feature of identity, because just as there are no two identical fingerprints, two same handwriting or two same retinas, there are no two identical voices.

Evidence based on voice analysis can be crucial for identifying, prosecuting and convicting the perpetrator. The information that is transmitted through speech communication is discrete in nature, because it consists of a series of elements from the set with a finite number of elements. The speed of the exchange of information, when it comes to speech, is determined by the physical limitations to which the human speech apparatus is subject. Forensic analysis of speech signal consists of three analyzes: auditory linguistic analysis, phonetic acoustic analysis and acoustic instrumental analysis.

Forensic analysis of the speech signal uses the achievements of several different sciences through the application of various complex methods. The multidisciplinary and complexity of the analysis makes this method of face identification extremely selective, and thus powerful.

**Key words:** forensic voice analysis, speech, evidence, crime.

## I. Introduction

Criminalistik as a science discipline deals with the prevention and suppression of crime with the help of practical methods and various means. As a scientific discipline, it is characterized by commitment to man as a complex being, his protection through the protection of health and rights of individuals, but also the protection of the health and rights of the community. The successful detection and investigation of criminal offenses depends on a series of actions and measures undertaken in the prevention and detection of perpetrators of criminal offenses through general and special methods<sup>5</sup>. Through the work of criminal experts to determine the truth and truthfulness, as well as the implementation of legality by judicial and competent authorities, the increase in crime is prevented. In this way, the coordination and efficiency of

---

<sup>1</sup> The work was created within the project of the Fund for Science of the Republic of Serbia "Ideje" - Project of acceleration of innovation and encouragement of entrepreneurship in the Republic of Serbia - Management of New Security Risks - Research and Simulation Development - NEWSIMR&D".

<sup>2</sup> Assistant, Faculty of Special Education and Rehabilitation, University of Belgrade, e-mail: [jadrankastevovic@yahoo.com](mailto:jadrankastevovic@yahoo.com), [orcid.org/0000-0001-7051-3897](https://orcid.org/0000-0001-7051-3897)

<sup>3</sup> Associate professor, University of Criminal Investigation and Police Studies in Belgrade, e-mail: [bozidar.otasevic@kpu.edu.rs](mailto:bozidar.otasevic@kpu.edu.rs), [orcid.org/0000-0002-6663-2874](https://orcid.org/0000-0002-6663-2874)

<sup>4</sup> Associate professor, Faculty of Law, University of Pristina, temporarily located in Kosovska Mitrovica, Serbia, e-mail: [sasa.atanasov@pr.ac.rs](mailto:sasa.atanasov@pr.ac.rs), [orcid.org/0000-0003-0157-3049](https://orcid.org/0000-0003-0157-3049)

<sup>5</sup> Modly, D., Šuperina, M., & Korajlić, N. (2008). Rječnik kriminalistike (in Croatian) (Dictionary of criminal science) Zagreb: Professional Association of Criminalists.

the authorities in the fight against all criminal acts and forms of criminality is achieved<sup>6</sup>. In modern criminal procedural practice, the existence of material evidence is of essential importance for the success of the procedure for determining the material truth. Forensics, as a multidisciplinary field, is tasked with providing material evidence of all kinds<sup>7</sup>.

One of the goals of forensics as a discipline is the identification of persons. Biological characteristics make man a unique being on the planet. It is this uniqueness that is the basis of the process of determining identity, that is, identification<sup>8</sup>. In this paper, we will talk about the human voice and the possibilities of identifying a person based on his voice. In everyday life, we identify and recognize the people around us by looking and/or listening. With this approach and analysis, it is not only important what is said, but also *how it is said*. It means that, not only the choice of words and expressions with which one wants to say something, but the way of pronouncing sounds, intonation, pace of speech, organization of utterances, loudness, etc., are important, too. In addition, it should be kept in mind that in the concept of how, the accompanying actions, whether we smiled, were angry, what was our body position, the direction of our gaze, etc., are hidden. This very fact shows that attention should be focused not only on what was said, but also on the way it was said, all with the aim of more complete and successful communication. During communication, in addition to the verbal part of the message transmission, there is also a non-verbal part. Both of these "parts" of communication are constantly present and intensively influence each other. For the purposes of this work, it is important to emphasize the non-verbal part of communication that is hidden in the term kinesics, which is the name for the systematic use of facial expressions to communicate meaning, especially in relation to the use of language (e.g. when frowning or speaking changes the interpretation of the sentence that has been spoken)<sup>9</sup>.

Each person's voice contains physiological and behavioral characteristics and is therefore very specific. It is the personal characteristic of each individual, such as a fingerprint, DNA profile, signature or eye structure. Biometric is a science that deals with examining those characteristics and features with the aim of confirming the identity of a person<sup>10</sup>. Voice analysis as one of the methods belongs to the biometric technique, which is based on the analysis of a person's behavior<sup>11</sup>. Biometric methods of voice analysis are mainly based on information obtained by analyzing acoustic features in the speech signal that are characteristic of each speaker. Forensic methods are based on the use of all possible individual markers (characteristics) that can be identified in the speech signal, as well as on the linguistic information that the speech signal carries. We can conclude that there is a significant difference in the analysis of the acoustic features of the speech signal for the needs of biometric and forensic methods. Namely, with the biometric method, the desire of the examinee is to be recognized, which means that the examinee is cooperative and that the final outcome is awaited, which is acceptance or non-acceptance by the system<sup>12</sup>.

With the forensic method, the goal of the interviewee (speaker) is to conceal the identity, the interviewee is uncooperative, and the final outcome may be a conviction for a criminal offense.

---

<sup>6</sup> Simonović, B. (2004). *Kriminalistika (in Serbian) (Criminalistic)*. Belgrade: Faculty of Law, Institute of Legal and Social Sciences.

<sup>7</sup> Otašević, J., & Otašević, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.

<sup>8</sup> Tuthill, H. (1994). *Individualization: Principles and procedures in criminalistics*, Salem: Lightning Powder Co.

<sup>9</sup> Kristal (1999) Enciklopedijski rečnik moderne lingvistike (in Serbian) (Encyclopedic Dictionary of Modern Linguistics). Belgrade: Nolit, 160.

<sup>10</sup> Ross, A, Jain, A, (2003) Information fusion in biometrics. *Pattern recognition letters*, 24(13), p.2118.

<sup>11</sup> Phillips, P.J., Martin, A., Wilson, C.L., Przybocki, M, (2000) An introduction evaluating biometric systems. *Computer*, 33(2), p.58.

<sup>12</sup> Phillips, P.J., Martin, A., Wilson, C.L., Przybocki, M, (2000) An introduction evaluating biometric systems. *Computer*, 33(2), p.60.

What also needs to be emphasized is the fact that with the forensic method, the content of the text being analyzed is beyond the control of the person doing the analysis, while with the biometric method, that content is carefully selected for the use of the biometric identification system<sup>13</sup>.

### **A Speech as a characteristic of the human beings**

Speech, viewed as a sound signal, carries a lot of information. In addition to the content of the spoken message, it also carries information about the speaker, his gender, age, origin, health and emotional state, level of education, etc. When understanding what is said, a person uses his knowledge of the language and understanding of the context<sup>14</sup>. That is why it is important to "listen and hear" in a special way, and then analyze. With this kind of targeted and purposeful listening, it is very important that the listening subject has a certain listening skill<sup>15</sup>. This is where the competencies and experiences of the forensic analyst come to the fore.

### **B Basic human voice generation**

Organs that directly participate in voice production are represented by the effector communicative system. The effector communicative system consists of four mechanisms: Initiation, phonation, oro-nasal process and articulation (respiration, phonation, resonance, and articulation). Each of these systems gives a characteristic to each speaker individually<sup>16</sup>. In addition to respiration, a person used the air current to generate voice, which circulates through the trachea-bronchial tree and lungs as the driving force for the vibrations of the vocal cords. Phonation requires a special control of expiration that is completely independent of the normal mechanism and rhythm of breathing. Optimal breathing with phonation is manifested in the fact that voice production is performed effortlessly, naturally and without tension. The larynx is the organ of phonation and voice generator<sup>17</sup>. It is an active cartilaginous tube, composed of several paired and unpaired cartilages<sup>18</sup>. The size of the larynx depends on the sex, age and individual characteristics of a person. The growth of the larynx leads to an increase in the vocal cords, which are 25 mm long in an adult man, and 18 to 20 mm in women<sup>19</sup>.

The voice resonator is the space where the basic laryngeal tone is amplified. In that space, resonance occurs, that is, when a medium under the influence of external vibration (vibrating energy) begins to vibrate itself. This is exactly what happens in this part of the voice generation process. A pure laryngeal voice is of very low intensity and without the participation of the resonator it would sound inhuman. The resonator must be matched to the tone source in size and shape, because there is an ideal resonator for each frequency. An ideal resonator must

---

<sup>13</sup> Otašević, J., & Otašević, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.

<sup>14</sup> Otašević, J., & Otašević, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.

<sup>15</sup> Otašević, J., & Otašević, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.

<sup>16</sup> Keramitičijevski, S.(1990) *Opšta logopedija (in Serbian) (General Speech and Language Therapy)*, Belgrade : Naučna knjiga.

<sup>17</sup> Cooper, M. (1977). Direct vocal rehabilitation. Approaches to vocal rehabilitation, ed. Cooper, M. and Cooper. *MH Springfield, III.: Charles C. Thomas.*

<sup>18</sup> The *plicae vocales* of the vocal cords are banded muscle folds that extend from the corner of the thyroid cartilage to the vocal process of the arytenoid cartilage. During the process of respiration, the glottis (opening between the vocal cords) has a triangular shape. During phonation, the vocal cords come close to each other and completely close the opening of the larynx. Their approach to the middle of the larynx and vibration leads to the creation of sound.

<sup>19</sup> Keramitičijevski, S.(1990) *Opšta logopedija (in Serbian) (General Speech and Language Therapy)*, Belgrade: Naučna knjiga.

amplify the fundamental tone and create the corresponding higher harmonic tones<sup>20</sup>. The resonator of the human voice is ideal and unique, because it has the ability to change the shape, volume and solidity of the walls. This allows the almost unlimited ability of this resonator to adapt to the source of the basic tone and to enable a whole range of variations that cannot be caused by any instrument constructed so far. The articulatory system by function participates in the generation of each voice (phoneme) individually. Each phoneme has its own acoustic features that are directly related to the position of the articulating organs that make up the articulating system<sup>21</sup>. It should be emphasized that speech, as a physical realization of language, in addition to the information it carries, also carries information about the speaker's psychological and emotional state<sup>22</sup>.

Laver (1980)<sup>23</sup> points out that an experienced expert can, on the basis of a certain person's voice and speech, evaluate his build and body size, gender, age, health condition, etc. The social position can also be determined in general, where a person comes from, as well as many other facts.

Acoustic features are characteristics of each segment of speech (voice, syllable, word, verbal expression) and these features are located within *variation fields*. The variation field represents a set of values of an acoustic characteristic that has the same perceptual meaning in the identification of a verbal event<sup>24</sup>. These variation fields represent "intra-speaker" and "inter-speaker" variations. Intra-speaker variations are consciously or unconsciously "imprinted" into the speech signal by each speaker. They are always present in spoken expression, and there are variations caused by emotions and conversation in specific situations (noisy environment - stadium, library, lecture...). Inter-speaker variations are characteristic of each speaker and they are markers of each individual (gender, age, dialect, accent, anatomical structure of the speech apparatus...)<sup>25</sup>. Variations in the speech signal (acoustic characteristics) are perceived and well recognized by the listener. It is these facts that make the human (listener) still irreplaceable in the process of identifying a person based on voice (speech). When it comes to speaker identification, when it comes to forensic requirements, elements of discrimination within variation fields are essential<sup>26</sup>.

### ***C Forensic voice and speech analysis***

In the forensic methods of voice and speech analysis, the following are most often used today: auditory linguistic analysis, phonetic acoustic analysis and hybrid forensic voice analysis (acoustic instrumental analysis)<sup>27</sup>. One of the features related to forensic methods is that they use all possible "markers" that can be registered in the speech signal, as well as the linguistic

---

<sup>20</sup> Heđever, M., Kovačić, G. (1997). *Acoustics of voice and speech*, Course script Speech acoustics for speech therapy students, Zagreb: Faculty of Education and Rehabilitation in Zagreb.

<sup>21</sup> Dobrota, N (2009). *Artikulationo fonološki poremećaji govora (Articulatory phonological disorders of speech)*, Belgrade: Institute for Psychophysiological Disorders and Speech Pathology "Cvetko Brajović" and Faculty of Special Education and Rehabilitation, University of Belgrade.

<sup>22</sup> Jovičić, S. T. (1999). *Speech communication: physiology, psychoacoustics and perception*, Belgrade : Nauka.

<sup>23</sup> Laver, J. (1980). The phonetic description of voice quality. *Cambridge Studies in Linguistics London*, 31, 1-186.

<sup>24</sup> Allen, J. B. (1994). How do humans process and recognize speech? *IEEE Transactions on speech and audio processing*, 2(4), 566.

<sup>25</sup> Antešević, S., & Jovičić, S.T. (2000). Značaj koartikulacionih i kontekstualnih elemenata u percepciji govora (Significance of co articulation and contextual elements in speech perception), *Nauka Tehnika Bezbednost*, 1-2.

<sup>26</sup> Jovičić, S.T. (2001) Forenzički aspekti prepoznavanja govornika (Forensic aspects of speaker identification), *Nauka Tehnika Bezbednost*, 1.

<sup>27</sup> Jovičić, S.T, Grozdić, Đ. (2014) Arguments for auditory-instrumental approach in forensic speaker recognition, *Proc. of Int. Sc. Conf. "Archibald Reiss Days"*, KPA, Belgrade.

information that each speech message carries<sup>28</sup>. In our country, there are no prescribed norms or protocols for such examinations, so the examinations are conducted in accordance with the code of the International Association for Forensic Phonetics (IAFP). Formulation of the final assessment is carried out according to the standards for comparing votes of the International Association for Identification adopted in 1991, which were expanded by the Forensic Science Service in 2000, based on the needs of practical experience<sup>29</sup>. The formulated scale of verbal ratings is used in most European countries, and it is also used here in the Laboratory for Forensic Acoustics and Phonetics, Center for the Improvement of Life Activities, and is shown in Table 1<sup>30</sup>.

Table 1: Vocal Identification Verbal Scale Grades

IAI standards (1991)	Forensic Science Service (2000)	Acoustic Phonetics and Forensics Laboratory Belgrade
Complete identification	Very strong likelihood	Very high degree of likelihood
-	Strong likelihood	High degree of likelihood
Probable identification	Moderately strong likelihood	Significant degree of likelihood
-	Moderate likelihood	Certain degree of likelihood
Possible identification	Limited likelihood	Likelihood exists and person cannot be excluded
Inconclusive case	Inconclusive case	Inconclusive case
Possible elimination	Limited unlikelihood	Unlikelihood exists and person cannot be excluded
-	Moderate unlikelihood	Certain degree of unlikelihood
Probable elimination	Moderately strong unlikelihood	Significant degree of unlikelihood
-	Strong unlikelihood	High degree of unlikelihood
Complete elimination	Very strong unlikelihood	Very high degree of unlikelihood

The essence of forensic recognition is the discrimination of characteristics in the speech of two speakers and the determination of criteria on the basis of which it can be determined with greater or lesser certainty whether their similarity exists. This assessment is given partially for each feature individually. The final evaluation is given by summarizing the analysis of all partial observations.

#### ***D Auditory linguistic and phonetic acoustic analysis***

These analyzes are carried out by listening to the recorded speech signals and observing certain linguistic characteristics of the speakers in the recording containing disputed and undisputed voice. What needs to be emphasized is that all these methodologies are applied after processing the spoken material (recording that exists and that has been submitted for expert opinion). This analysis is based on the auditory experience of the forensic expert, on his subjective observation and assessment. It includes analysis of voice quality, specific phonetic-linguistic manifestations

<sup>28</sup> Kašić, Z. & Đorđević, J. (2009) About linguistic expertise in crime solving. Proceeding of First conference with international participation: *Law and forensics in criminal law*, p. 333-340, Belgrade: Academy of Criminalistic and Police Studies.

<sup>29</sup> Bijhold, J., Ruijrok, A., Jessen, M., Geradts, Z., Ehrhardt, S., & Alberink, I. (2007, October). Forensic audio and visual evidence 2004-2007: A review. In *15th INTERPOL forensic science symposium*.

<sup>30</sup> Jovičić, S.T. (2001) Forenzički aspekti prepoznavanja govornika (Forensic aspects of speaker identification), *Nauka Tehnika Bezbednost*, 1.

(which were discussed in the introductory part), changes in vocal expression, pathological changes in sounds and analysis of psychological and emotional state. It states the specific features of individual sounds, groups of sounds in disputed and undisputed voice. It is precisely in this part of the assessment that the spectrogram results (when the data obtained by applying both analyzes are crossed and combined) are of great help. It is known that the speaker's psychological and emotional state greatly influences changes in acoustic features in the time, frequency and intensity domains (prosody), while it has very little influence in the formant domain. In addition, this analysis evaluates several aspects of language, namely phonological-phonetic, dialectological, psycholinguistic and sociolinguistic.

### **E Hybrid forensic voice analysis**

Acoustic instrumental analysis represents the application of methods using modern technology (computers) in order to obtain certain data contained in the acoustic wave, which cannot be obtained using only the sense of hearing<sup>31</sup>. In the context of speech signal processing, the acoustic theory of speech has proven to be useful for practical application. In order to analyze the speech signal with a computer, the acoustic wave needs to be converted into an electrical signal. It is the microphone that converts the acoustic wave into an electrical signal, but in the sound cards that almost all computers have, this signal is additionally adapted to multiple applications<sup>32</sup>. Then, many complicated methods are applied to that received signal, which gives the results of discrimination in disputed and undisputed voice. The methods used are: spectrogram, Fourier transformations, DFT (Discrete Fourier Transform) and FFT (Fast Fourier Transform), cepstrum analysis, analysis by means of spectral moments, LR – *Likelihood Ratio* and others.

### **F Spectrogram**

Spectrogram testing involves visualizing the speech signal on a screen or with special instruments – sound analyzers. The result of the examination is a spectrogram – an image of a person's voice (voice print). In this way, the perpetrator of the crime could be identified by comparing the sound image – spectrogram, of the perpetrator with the sound images from the database. However, this is not possible because the human voice is a variable quantity, even when it belongs to the same person, the voice is prone to change due to a number of factors. Therefore, the result of the obtained identification is expressed only by a certain degree of probability of reliability of identification or elimination.

The basic acoustic characteristics of phonemes can be observed by analyzing the spectrogram. Each language has a different number of phonemes (between 30 and 50). The English language has 42 phonemes (these are graphemes that appear in the mark for reading English words), while in the Serbian language each phoneme corresponds to one letter of the alphabet and there are 30 of them<sup>33</sup>. The speed of information exchange, when it comes to speech, is determined by the physical limitations of the human speech apparatus. These limitations are such that in normal speech a person produces an average of 10 phonemes per second<sup>34</sup>. The spectrogram shows the speech signal in three dimensions. It shows the duration of the speech signal in seconds (s), the frequency in Hertz (Hz) and different shades of gray, black and white, which

---

<sup>31</sup> Nešić, L., Kovačević, J., Stevović-Otašević, J. (2011). Forensic analysis of the speech signal (Forenzička analiza govornog signala - in Serbian), *Pravni život*, 60 (5-6), 90.

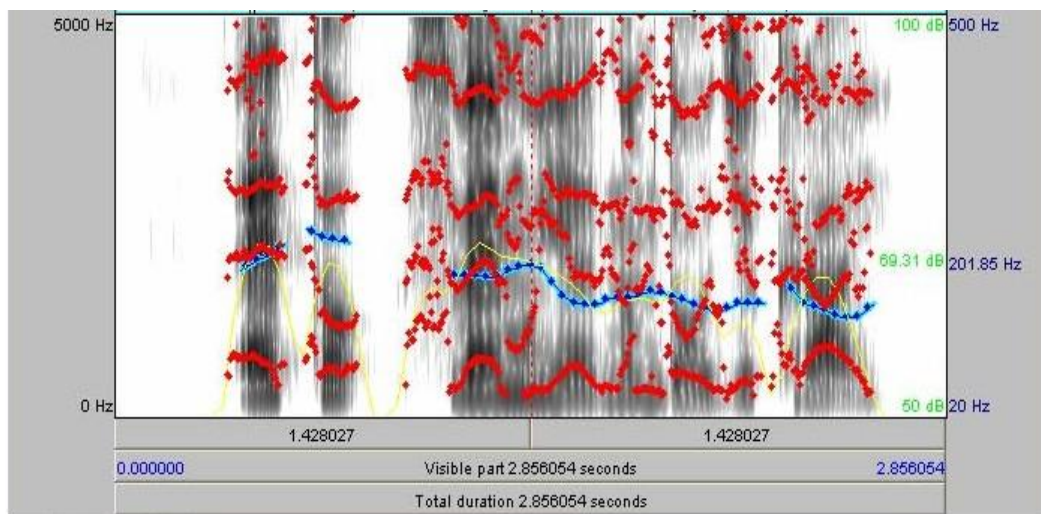
<sup>32</sup> Nešić, L., Kovačević, J., Stevović-Otašević, J. (2011). Forensic analysis of the speech signal (Forenzička analiza govornog signala - in Serbian), *Pravni život*, 60 (5-6), 93.

<sup>33</sup> Nešić, L., Kovačević, J., Stevović-Otašević, J. (2011). Forensic analysis of the speech signal (Forenzička analiza govornog signala - in Serbian), *Pravni život*, 60 (5-6), 87 – 102.

<sup>34</sup> Otasevic, J., & Otasevic, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.

represent a measure of the strength of the frequency components of the speech signal in a given unit of time. The white color indicates the existence of the strongest frequency, it is the most pronounced (it is characteristic of vowels), while shades of gray and black show noisy parts of the spectrum, characteristic of all consonants. Also on the spectrogram you can see the formants of the human voice, which represent spectral areas (peaks) and are characteristic of vowels as well as the consonants M and N<sup>35</sup>. The frequency range of formants ranges from 250 Hz to 2.5 kHz<sup>36</sup>. They represent the individual characteristics of each speaker and are extremely important in the identification process because they can distinguish or recognize people. Differences in formants are also influenced by coarticulation of adjacent phonemes in a word, as well as accent.

Picture 1: Spectrogram of speech signal



A more recent application of forensics in the field of voice and speech is based on the results obtained from auditory analysis and acoustic analysis that uses a *Bayesian* model to assess the reliability of evidence<sup>37</sup> based on the assessment of the probability of two opposing assumptions. And that is the prosecutor's assumption that the disputed voice/speech belongs to the suspect, and the second assumption is that it belongs to someone else from the possible suspect population. The relationship between these two assumptions is called the likelihood ratio (LR – *Likelihood Ratio*) and is expressed as a numerical value. This value shows the relative strength of the evidence and is an additional element to the prosecution, which makes their position stronger or weaker.

## II. Conclusion

Forensic analysis in practice is performed on a large number of features (characteristics) that can be separated and analyzed in the available recordings. The essence of forensic processing of voice and speech is the application of various methods, which obtain a large number of

<sup>35</sup> Formants are indicators of the resonance of the vocal tract during the speech process. These are the frequencies of the strongest air current that passes through the vocal tract.

<sup>36</sup> Bijhold, J., Ruijck, A., Jessen, M., Geradts, Z., Ehrhardt, S., & Alberink, I. (2007, October). Forensic audio and visual evidence 2004-2007: A review. In *15th INTERPOL forensic science symposium*.

<sup>37</sup> Bakır, Ç., & Yüzkat, M. (2022, May). A Search on the Importance of Forensic Voice Studies in Forensic and a Example Application. In *2022 30th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.

results, and which in a certain way should suggest the final assessment and position in the comparison of two voices (disputed and undisputed). With this procedure, speaker recognition takes on the character of a multidimensional space and a multidisciplinary approach. The recordings that are analyzed are of different quality and often some features are impossible to identify, in addition, not all the features that are evaluated have the same importance and weight for assessing the individuality of the voice/speech. All these facts indicate the seriousness of the assessment procedure and require the work of experts in these fields. Material evidence is most often presented by experts, and the statements of witnesses and the suspect/defendant are confirmed or denied. Only evidence whose authenticity and evidentiary information derived from it is confirmed by experts and/or witnesses or/and suspects/defendants has probative value.

### III. References

- Allen, J. B. (1994). How do humans process and recognize speech? *IEEE Transactions on speech and audio processing*, 2(4), 567-577.
- Antešević, S., & Jovičić, S.T. (2000). Značaj koartikulacionih i kontekstualnih elemenata u percepciji govora (Significance of co articulation and contextual elements in speech perception-in Serbian), *Nauka Tehnika Bezbednost*, 1-2.
- Bakır, Ç., & Yüzkat, M. (2022, May). A Search on the Importance of Forensic Voice Studies in Forensic and a Example Application. In *2022 30th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). IEEE.
- Bijhold, J., Ruifrok, A., Jessen, M., Geradts, Z., Ehrhardt, S., & Alberink, I. (2007, October). Forensic audio and visual evidence 2004-2007: A review. In *15th INTERPOL forensic science symposium*.
- Cooper, M. (1977). Direct vocal rehabilitation. Approaches to vocal rehabilitation, ed. Cooper, M. and Cooper. *MH Springfield, III.: Charles C. Thomas*.
- Dobrota, N (2009). *Artikulaciono fonološki poremećaji govora (Articulatory phonological disorders of speech)*, Belgrade: Institute for Psychophysiological Disorders and Speech Pathology "Cvetko Brajović" and Faculty of Special Education and Rehabilitation, University of Belgrade.
- Heđever, M., Kovačić, G. (1997). *Akustika glasa i govora (Acoustics of voice and speech)*, Course script Speech acoustics for speech therapy students, Zagreb: Faculty of Education and Rehabilitation in Zagreb.
- Jovičić, S. T. (1999). *Govorna komunikacija - fiziologija, psihoakustika i percepcija (Speech communication: physiology, psychoacoustics and perception)*, Belgrade : Nauka.
- Jovičić, S.T, Grozdić, Đ. (2014) Arguments for auditory-instrumental approach in forensic speaker recognition, *Proc. of Int. Sc. Conf. "Archibald Reiss Days"*, KPA, Belgrade.
- Jovičić, S.T. (2001) Forenzički aspekti prepoznavanja govornika (Forensic aspects of speaker identification), *Nauka Tehnika Bezbednost*, 1.
- Kašić, Z. & Đorđević, J. (2009) About linguistic expertise in crime solving. Proceeding of First conference with international participation: *Law and forensics in criminal law*, (pp. 333-340) Belgrade: Academy of Criminalistic and Police Studies.
- Keramičijevski, S.(1990) *Opšta logopedija (in Serbian) General Speech and Language Therapy* Belgrade : Naučna knjiga
- Kristal (1999): *Enciklopedijski rečnik moderne lingvistike (in Serbian) Encyclopedic Dictionary of Modern Linguistics*. Belgrade: Nolit
- Laver, J. (1980). The phonetic description of voice quality. *Cambridge Studies in Linguistics London*, 31, 1-186.



- Modly, D., Šuperina, M., & Korajlić, N. (2008). Dictionary of criminal science (Rječnik kriminalistike - in Croatian) *Zagreb: Professional Association of Criminalists*.
- Nešić, L., Kovačević, J., Stevović-Otašević, J. (2011). Forenzička analiza govornog signala (in Serbian) Forensic analysis of the speech signal, *Pravni život*, 60 (5-6), 87 – 102.
- Otasevic, J., & Otasevic, B. (2021). Voice-Based Identification and Contribution to the Efficiency of Criminal Proceedings. *J. Crimin. & Crim. L.*, 59, 61.
- Phillips, P.J., Martin, A., Wilson, C.L., Przybocki, M, (2000) An introduction evaluating biometric systems. *Computer*, 33(2), pp.56-63
- Ross, A, Jain, A, (2003) Information fusion in biometrics. *Pattern recognition letters*, 24(13), pp.2115-2125.
- Simonović, B. (2004). *Kriminalistika (Criminalistic) (in Serbian)*.Belgrade: Faculty of Law, Institute of Legal and Social Sciences.
- Tuthill, H. (1994). *Individualization: Principles and procedures in criminalistics*, Salem: Lightning Powder Co.