ARTICLE

Forensic Linguistics: Identification of Individuals by Written and Oral Speech as Evidence in Criminal Cases in BRICS Countries (Brazil, India, Russia)

Julia Morozova,

South Ural State University (National Research University) (Chelyabinsk, Russian Federation)

https://orcid.org/0000-0001-9361-4703

Natalie Papoyan,

South Ural State University (National Research University) (Chelyabinsk, Russian Federation)

https://orcid.org/0000-0001-6545-2962

Niteesh Kumar Upadhyay,

Tashkent State Law University (Tashkent, Uzbekistan)

https://orcid.org/0000-0003-3063-2431

https://doi.org/10.21684/2412-2343-2025-12-4-119-146

Received: March 10, 2025
Reviewed: April 22, 2025
Accepted: October 1, 2025

Abstract. This paper explores the dimensions for systematic collaboration in forensic linguistics among BRICS nations, specifically Brazil, India, and Russia. In an increasingly globalized world where crimes are frequently committed through text or speech, the topic is vital. A review of scientific literature reveals that while some countries possess solid domestic foundations in forensic linguistics–encompassing forensic phonetics and authorship analysis–the focus has remained on national applications, overlooking opportunities for exchanging knowledge internationally. The research employs a "compare and contrast" methodology to systematically analyze

the techniques, legal frameworks, and applications in these three countries. Russia's approach is characterized by scientifically grounded techniques using phonetics, acoustics, and advanced quantitative software. Brazil similarly employs precise perceptual-auditory, acoustic, and stylometric analyses. In contrast, Indian experts have recognized the need to integrate such methodologies into their justice system. The article argues that mutual exchange of linguistic knowledge and practical experience can refine existing methodologies, help establishing effective practices in India, and ultimately enhance crime investigation and resolution across the BRICS+ alliance. Such collaboration is particularly crucial for combating anonymous online crimes, where advanced linguistic analysis can significantly improve investigative efficiency.

Keywords: forensic linguistics; authorship attribution; analysis; speaker recognition; criminal proceedings, BRICS.

To cite: Morozova, J., Papoyan, N., & Upadhyay, N. K. (2025). Forensic linguistics: Identification of individuals by written and oral speech as evidence in criminal cases in BRICS countries (Brazil, India, Russia). *BRICS Law Journal*, *12*(4), 119–146.

Table of Contents

Introduction

All human thoughts are expressed through language; therefore, scholars have always focused on the role of language in various aspects of human life. The study of the relationship between language and law is also highly relevant. On the one hand, all laws are formulated in words, and there is no law without language. On the other

hand, language can be used as a tool for committing crimes or transmitting evidentiary information. Thus, linguistic knowledge is actively employed in legal sciences.

The potential of this area is becoming increasingly vital in today's globalized world, where crimes are often committed through text or speech. A review of scientific literature and expert practices in the BRICS countries–Brazil, India, and Russia–reveals a solid foundation of forensic linguistics. Despite significant national developments, the current literature focuses on domestic applications and methodological improvements. However, one critical aspect has been overlooked: the potential for systematic knowledge exchange and collaboration in forensic linguistics.

Therefore, this paper aims to explore the possibilities for mutual exchange of linguistic knowledge and practical experience in order to enhance crime investigation and resolution across three countries. The core of the research relies on the comparative-contrastive method, which systematically compares and contrasts the techniques, legal frameworks, and applications used in Brazil, India and Russia. The authors seek not only to catalog differences but also to identify synergies that can contribute to the advancement of forensic linguistics within the BRICS alliance and beyond, and to identify areas where collaboration and knowledge exchange can be beneficial.

Forensic linguistics is a relatively new but actively developing area of applied linguistics. It involves the application of linguistic methods and knowledge to resolve legal issues. The primary tasks of forensic linguistics include text analysis, identification of the author's individual features, assessment of the accuracy and reliability of translations, and interpretation of speech data within the context of legal proceedings. Linguists provide expert opinion in the investigation of crimes related to speech offenses. This helps investigators and courts to make well-founded decisions.

The role of linguists in the legal sphere is becoming increasingly significant. The integration of linguistics and law is a crucial step in modern legal practice. Coulthard, Johnson and Wright state,

> Since 1993 forensic linguistics has been gaining ground as a very broad and diversified field of applied speech science all over the world.[1]

Forensic linguistics was initially used to obtain evidence in crime investigations. However, as it developed, forensic linguistics came to be used in its generally accepted broader meaning. The main global areas of study and application of forensic linguistics are the written language of the law, particularly the language of legislation; spoken legal discourse, particularly the language of court proceedings and police questioning; the social justice issues that emerge from the written and spoken language of the law; the provision of linguistic evidence, which can be divided into evidence on identity/authorship and evidence on communication; the teaching and learning of spoken and written legal language; and legal translation and interpreting.[2]

---

[1] Coulthard, M., Johnson, A., & Wright, D. (2016). *An introduction to forensic linguistics: Language in evidence*. Routledge.

[2] Gibbons, J., & Turell, M. T. (Eds.) (2008). *Dimensions of forensic linguistics*. Equinox Publishing.

The studies aimed at identifying the criminal are essential for investigating and solving criminal offenses. In this area, forensic linguistics is represented by two lines: forensic phonetics and authorship studies. This paper considers the general principles of expert research in such BRICS countries as Brazil, India and Russia.

1. Meaning of the Term "Forensic Linguistics"

In Russia, linguists have been involved in court proceedings since the end of 1991. This period was characterized by the emergence of "glasnost" (government transparency) and the need to find a line between ensuring freedom of speech and protecting human rights. Forensic linguistic expertise appeared a little later, in the 2000s. It is defined as

> a multidisciplinary field in the theory of linguistic expert science, combining criminalistics and linguistics, investigation of the crime scene and speech examination, forensics and discourse analysis, criminology and the theory of speech acts, the theory of expertise and textology.[3]

The term "Legal Linguistics" is used in Russia to define a field of science studying the interaction of language and law. This term was introduced into scholarly discourse by the famous Russian linguist N.D. Golev, Doctor of Philology, Professor of Kemerovo State University. The implementation of "*Yurislingvistika*" (*Legal Linguistics*), a unique scientific project and journal, contributed to the development of legal linguistics in Russia. In their publications, scholars search for new tools and methods for studying legal communications, legal writing, linguistic expertise, language policy, and legal education.[4] Besides, the interpretation of legal texts, the distinction between legal and legislative techniques, the linguistic analysis of legal documents, the development of recommendations for creating texts of laws and other regulatory instruments, studies into legal translation, as well as studies into legal, psychological, and social aspects of speech disputes are relevant for the Russian legal system.

Higher education in forensic linguistics is highly developed in Russia. In Russian universities, students can study subjects such as "Speech Expertise," "Analysis of Speech Activity Products,""Forensic Linguistic Expertise," and "Jurislinguistics," among others.

In the Russian Federation, forensic linguists are engaged in various governmental organizations: the Ministry of Internal Affairs, the Ministry of Justice, the Federal Security Service, etc. Criminal expertise centres in these organisations have a large team of experts, special linguistic research methods, and unique software. Linguists working in Russian state expert institutions perform the following types of expertise:

---

[3] Galyashina, E. I. (2018). Forensic linguistics in Russia: State of the art and new challenges. *Theory and Practice of Forensic Examination*, *13*(4), 28–37. (In Russian).

[4] Golev, N. D. (1999). The legal aspect of language in linguistic coverage. *Legal Linguistics*, *1*(1), 12–59. (In Russian).

phonoscopic (Forensic phonetic expertise), linguistic, and authorship studies. In case the speaker or the author of a written text should be identified, law enforcement officers and judges usually turn to state expert organizations when investigating criminal cases. However, linguistic research can be carried out in private expert organizations. Private experts take on part of the work of state experts, while the defense can obtain an alternative opinion on linguistic expertise issues, which contributes to the observance of the adversarial principle in criminal proceedings.

In Brazil, forensic linguistics as a line of applied linguistics is divided into three branches: the analysis of legal language, the analysis of interaction in legal contexts (including translation and interpretation), and the study of language as evidence. For most experts, the term "forensic linguistics" is strictly defined and mainly identified with the last of the three mentioned areas. For others, it is broader and equated with the term "language and law." This difference in meanings can be seen in the names of associations organized by experts in this field. For example, the association of Portuguese-speaking forensic linguists is called ALIDI: Associação de Linguagem e Direito (although the practical application of linguistic analysis in forensic contexts in Portuguese-speaking countries is still minimal), and the worldwide association is called IAFL: International Association of Forensic Linguists.[5]

Scientific works on forensic linguistics in Brazil address issues such as simplifying the language of laws and regulations, as well as interpreting individual legislative instruments to ensure access to justice for all segments of the population. Interest in so-called "speech crimes" (in the source, "*crimes de linguagem*") has also increased, as evidenced by the case in which a judge of the Supreme Court of Brazil, Gilmar Mendes, filed a lawsuit against Monica Iozzi, a Brazilian journalist and actress, claiming that he was the victim of an insult to his honour. Iozzi posted an image of the judge on her Instagram account captioned "Accomplice?" Translations for foreign witnesses, victims, and suspects remain relevant. Additionally, linguists provide lectures to police officers and lawyers on the interrogation of witnesses and defendants, as well as on behavior toward individuals who do not speak the official language of the legal system.

Brazilian law and legal science focus on the problem of violence against women. Forensic linguistics studies legal discourse in rape cases. An analysis of court rulings on cases related to this crime shows that the formulations of judges do not always clearly indicate the full responsibility of the rapist for his actions. These texts often present the victim as someone who shares part (if not all) of the blame for the incident.[6]

Rui Sousa-Silva highlights the significant development of forensic linguistics in the context of the Portuguese language over the last decade. Firstly, the Language and

---

5 Almeida, D. C. D. (2015). *Análise forense de autoria textual: Estilos sociais e individuais* (Doctoral dissertation). Universidade de São Paulo. (In Portuguese).

6 Figueiredo, D. (2014). Discurso, gênero e violência: uma análise de representações públicas do crime de estupro. *Language and Law/Linguagem e Direito*, *1*(1), 141–158. (In Portuguese).

Law Association (ALIDI) was founded in 2014 to bring together linguists and lawyers working on the relationship between language and law. Secondly, the Language and Law/Linguagem e Direito international bilingual journal was set up in the same year as a result of a partnership between the Federal University of Santa Catarina (UFSC) and the University of Porto. This journal presents scientific papers touching upon current law and forensic linguistics in Brazil. Thirdly, it is worth mentioning the educational activities of federal universities in forensic linguistics. They offer not only master's and doctoral studies in "Linguistics," but also have a specialised course in forensic linguistics supported by the Portuguese Association of Forensic Sciences. Rui Sousa-Silva notes the prospects for the development of this field in Portuguese-speaking countries, which requires an increase in the number of scientific studies and publications based on a rigorous scientific approach and interaction with law enforcement agencies.[7] Leonardo Vichi indicates that forensic linguistics in Brazil requires its own individual approach and development, since "English-language ethnocentrism predominates in the literature on this issue, which does not take into account the peculiarities inherent in the Portuguese language and complicates the work of professional native speakers of Brazilian Portuguese."[8]

In Brazilian criminal proceedings, forensic phonetics and forensic authorship expertise are used to identify individuals by written or spoken language.

Forensic expert activities in Brazil are carried out by the Polícia Científica, which is a public administration body operating in most Brazilian states. The Scientific Police reports directly to the Secretariats of Public Security or similar bodies, except for some states, where it remains part of the civil police structure and closely cooperates with the civil and military police.

In India, characterised by multilingualism and dialectal diversity, forensic linguistics offers high potential for investigations, trials, and interpretation of laws.[9] Language interacts with law under the Official Languages Act adopted in 1963. The Department of State Language was established in 1975 to increase annually the volume of official correspondence carried out in Hindi in the Devanagari script. The Parliamentary Committee on the State Language has set a goal of using Hindi in all communications by the national Independence Day of 2047, when the country will celebrate 100 years of freedom. Thus, the transition to Hindi is being carried out stepwise by adding more than a thousand new words from other local languages. Forensic linguistics in its broad sense can certainly update and improve criminal investigation in India, especially given the large number of Hindi dialects, which often

---

[7] Gonçalves, J. S. S. (2021). Práticas de Análise em Linguística Forense. *Language and Law/Linguagem e Direito*, *8*(1), 120–124. (In Portuguese).

[8] Vichi, L. (2023). *O Perfilamento Criminal através da Linguística Forense Computacional*. Leonardo Vichi. https://leonardovichi.com/o-perfilamento-criminal-atraves-da-linguistica-forense-computacional/

[9] Prema, S. (2018). *A note on forensic linguistics in India*. Academia.edu. https://www.academia.edu/36172185/A_Note_on_Forensic_Linguistics_in_India

complicate the parties' understanding of criminal proceedings. For example, if the investigator records the statement of the accused in the investigator's dialect, it may distort the essence of what was said. Section 277 of the Indian Criminal Procedure Code, 1973, ensures that trials are conducted in the language of the accused to prevent misrepresentation of information and ensure a fair trial, thereby ensuring justice. These factors make forensic linguistics essential for the justice system in India. Notably, artificial intelligence technologies are being introduced in India to improve the efficiency of legal proceedings. Nomology Technology's Trees natural language processing platform helps transcribe court proceedings. The Supreme Court's AI portal, SUPACE, uses machine learning for data management. Artificial intelligence systems, such as SUVAAS, facilitate rapid translation of court rulings from English into local languages. Unfortunately, forensic linguistics in India is yet to receive the appropriate attention it deserves. Regrettably, India is lagging behind in terms of progress in this area. Forensic linguistic intervention in criminal investigations, trials, and the interpretation of law should be made mandatory in India.

An analysis of the scientific literature and practical activities of forensic experts in the countries under study demonstrates certain similarities in the activities of experts in Russia and Brazil. For example, the relationship between the roles of a state expert and a specialist engaged by the defence is perceived similarly. According to João Pedro Pádua,

> The main expertise in a case does not "belong" to the parties and is not promoted by them: it belongs to the court and is carried out by official expert bodies (most often in criminal proceedings) or experts the judge trusts (most often in civil proceedings).[10]

The law also distinguishes between the rights and obligations of a state expert and a specialist in the Russian Federation and an expert and a technical assistant (*assistentes técnicos*) in Brazil. The system logic suggests that only a state expert is presumed to be impartial. In fact, experts hired by the defence have already entered the process under the judge's suspicion. There are empirical data on the mistrust of judges in the evidence presented by the defense.[11] In the Russian Federation, expert opinion is becoming increasingly important in the investigation of criminal cases because of the need to comply with the adversarial principle. In Brazil, technical assistants are prohibited from drafting reports that could compete with the official expert report. They can only express their opinion on the report. The opinion may be critical, but it lacks the same force as an expert opinion and cannot replace it.

Forensic linguistics is an actively developing area of legal science. The rapid increase in the volume of text and speech information on the Internet necessitates

---

[10] Prema, 2018, p. 5.

[11] Lupetti Baptista, B. (2012). *Entre "quereres" e "poderes": Paradoxos e ambiguidades na imparcialidade judicial* (Doctoral thesis presented to the Graduate Program in Law). Universidade Gama Filho. (In Portuguese).

legal regulation of this area. In Russia and Brazil, forensic linguistics is understood in a broad and narrow sense. In India, the use of linguistic knowledge in the judicial system has not yet become a widespread practice.

2. Forensic Phonetics

Oral speech is the most common means of communication. There are many modern communication channels used for comfortable communication. However, criminals can use these channels to commit unlawful acts. For example, an attacker can call the police and report a bomb planted in an airport building. As a result, emergency services, including firefighters, ambulances, and police, are forced to arrive at the scene to evacuate people quickly. This leads to flight delays, which entail significant financial losses for airlines. In most cases, such reports turn out to be false. Identifying a criminal based on the audio recording of a telephone conversation is often the only evidence of the suspect's guilt. In such situations, state-of-the-art technologies and methods of analyzing audio recordings help law enforcement agencies. Phonetic experts compare the recording of a telephone call with the speech samples of the suspect.

Forensic speaker identification–also known as forensic voice comparison, voice recognition, or speaker comparison–is carried out for the purposes of assisting the police and courts to form a view concerning the identity of a person speaking in a speech recording. Specifically, speaker recognition is defined as the process of identifying an unknown speaker from samples of their voice.[12] The criminal's "disputed" or "questioned" speech samples may be recordings of telephone calls or covert recordings made by police or witnesses (for example, in cases involving bribery, extortion, or blackmail). These criminal recordings are then compared against known recordings of the suspect's speech, which most often consist of recorded police interviews.[13] So the main question is whether the disputed audio recording contains the suspect's voice and speech. Phonetics is concerned with how and where speech sounds are produced (articulatory phonetics); how speech sounds are perceived (auditory phonetics); and speech as wave structure (acoustic/experimental phonetics).[14]

In different countries, experts use different methods of speaker recognition from audio recordings. There are two main traditions of analyzing and comparing disputed audio recordings and speech samples: auditory perception and acoustic analysis. Acoustic analysis involves the use of specialised computer hardware and software to evaluate quantitative indicators and measure the parameters of oral speech.[15]

---

[12] Hollien, H. (1990). *The acoustics of crime: The new science of forensic phonetics*. Springer.

[13] Hollien, 1990, p. 1.

[14] Olsson, J., & Luchjenbroers, J. (2013). *Forensic linguistics.* Bloomsbury.

[15] Rusman, G., D'Orio, E., Popova, E., & Kipouras, P. (2023). Features of the application of digital technology in criminal proceedings of the BRICS countries. *BRICS Law Journal*, *10*(1), 35–58.

The "voiceprint" method was very popular a few decades ago. As the name suggests, a voiceprint is an analogue of a fingerprint. This method was developed by Lawrence Kersta in the United States in the early 1960s, based on spectrogram analysis, and promised 99% speaker recognition accuracy. Although this method was widely disputed, it was adopted in many court cases in the United States, and many police officers were trained to conduct the analysis. The reliability of this method has been debated for several reasons. Firstly, fingerprints are static, while speech is dynamic, which makes the comparison between them controversial. Secondly, the results of real experiments showed a significantly lower level of method accuracy. Thirdly, there was no theoretical basis for this methodology. However, according to Harry Hollien, "Voiceprints are a problem that simply will not go away."[16] In the 1970s, in response to the criticism of the voiceprint method, it was supplemented by the listening procedure, which became known as the aural-spectrographic method. However, this did not satisfy critics, since, according to P. Rose,

> the bad reputation of the aural-spectrographic method is mostly because of its lack of an explicit theoretical base.[17]

In other countries, the auditory recognition of the speaker prevails. This method involves forensic phoneticians listening to the audio recordings of speech and making a detailed phonetic transcription using the signs of the International Phonetic Alphabet (IPA). Thus, they identify the features of the speaker's oral speech and compare them with the features of the suspect's oral speech. This requires a highly qualified expert with both specialized education and experience, as well as a sensitive ear.

Another area in speaker recognition by oral speech is automatic speaker recognition using specialized software. The rapid development of speech technologies enables the reliable verification of the speaker in certain situations, such as when a user gains access to a mobile device, premises, or a bank account, using voice as biometric data. However, these situations are incomparable with the tasks of forensic phonetics. Reliable automatic speaker recognition can only be achieved with ideal audio recording quality (signal-to-noise ratio, frequency range, distortions caused by the uneven frequency response of the recording channel), duration and comparability of signal characteristics, which is unattainable in criminal proceedings.

2.1. Historical Sketch of Forensic Phonetics in BRICS Jurisdictions

Let us compare the development and establishment of forensic phonetics in three BRICS countries: Brazil, India, and Russia, starting with forensic phonetics in Russia, which has a history of over 50 years. Speaker recognition was first used in the summer of 1949 during a manhunt in the USSR. It was an investigation into the case of a Soviet diplomat who called the U.S. Embassy and tried to sell information about Soviet

[16] Hollien, 1990, p. 12.

[17] Rose, P. (2002). *Forensic speaker identification*. CRC Press.

foreign agents. Since then, active work on speaker recognition by oral speech has continued in the Russian Federation. Now forensic phonetics in the Russian Federation has a solid scientific base and wide practical application in criminal proceedings.

In Brazil, forensic phonetics is the most institutionalised and popular aspect of forensic linguistics. The Magri case was the first in the history of forensic phonetics in Brazil. In 1992, the Minister of Labor and Social Security, Antonio Rogerio Magri, admitted to the Director of the Collection Department that he had accepted a $30,000 bribe to help release funds to a construction company. The National Institute of Forensic Science had to enlist the help of doctors from UNICAMP, the State University of Campinas, to verify whether the speech on the recording belonged to the Minister and the Director. Since then, the National Institute of Forensic Science has invested in equipment and training for experts able to identify individuals from audio recordings. The first seminar on forensic phonetics was organised in 1994.[18]

There are more than 2,000 dialects in India, and more than 22 languages are recognized as official under the Indian Constitution. Each dialect has its own features depending on the geographical region. The speech patterns of Indians can provide valuable insights into the speaker's social status, place of birth, current location within the country, and educational background. In 2014, Nasib Singh, a police constable, was accused of demanding and accepting bribes. The complainant and his friend set up a sting operation in which the accused was seen asking for a bribe. The Central Bureau of Investigation sent the hard drives and memory cards to the CFSL in New Delhi for an opinion on whether they had been tampered with. The experts concluded that the memory card had not been altered or tampered with. Witnesses and complainant Chetan Sharma identified the accused's voice in the audio-video transcription of the recorded conversations when the voice parade was conducted, and the Inspector and colleagues who had worked with the accused in the same police station also identified the accused's voice. The voice sample of the accused could not be recorded as he refused to provide it. At the same time, the complainant's and witnesses' voices were recorded and sent to the CFSL for voice comparison. An expert witness from the CFSL in New Delhi conducted a comparative analysis of the voices of the complainant and witness in the video. They then compared these voices with a specimen voice and concluded that the voices in the recording and the specimen voice were similar and belonged to the same person.[19] After that, however, the development of forensic phonetics in India was somewhat limited. At present, there is no adequate scientific basis for voice and speech identification in India.

---

[18] Mattos, J. (2008). *Um estudo comparativo entre o sinal eletroglotográfico e o sinal de voz* (Dissertation submitted to the Graduate Program in Telecommunications Engineering). Fluminense Federal University. (In Portuguese).

[19] Toppo, R., & Sinha, S. (2023). Evaluating and accessing the scope of forensic linguistics in a multilingual context in India. *Language in India*, *23*(8), 113–159.

2.2. Categories of Criminal Cases

The example above exemplify the use of the suspects' voice and speech recognition as the main evidence of guilt in cases of false bomb threats. Experts utilize specialized digital sound processing programs to accentuate key elements of speech, amplify weak signals, and eliminate extraneous noise. These technologies significantly increase the chances of successfully identifying a criminal.

However, criminals are also utilising digital advances to conceal their identities. Some of them use special voice-altering software, which significantly complicates identification. In such cases, a multidisciplinary approach becomes essential, involving phonoscopists, linguists, and IT and cybersecurity specialists in the analysis.

Voice phishing (also known as "vishing") is another example of voice crime that has become increasingly common. Although voice phishing is a relatively new crime, it has become one of the most serious problems in recent years. It poses a significant threat to the financial well-being of citizens, particularly vulnerable groups. It is a shining example of how criminals are effectively using advanced developments in voice technology. Unfortunately, Russia, Brazil, and India are among the top ten countries whose citizens are most vulnerable to voice phishing attacks via WhatsApp Messenger. Speaker recognition by oral speech is an effective tool in investigating such crimes.

Speech as a method of committing a crime is also used in more serious cases, such as blackmail, death threats, premeditated murder, and kidnapping. Additionally, investigative bodies and judges effectively utilize forensic phonetic examination as evidence in cases involving embezzlement, economic fraud, corruption, and drug trafficking, among others. Forensic phonetic examination enables the identification of a criminal's voice, which can be used as conclusive evidence in investigations and court proceedings. Precise voice analysis techniques, the use of spectrograms, and other modern technologies help to identify unique features of the voice apparatus that cannot be faked or altered.

Many high-profile cases involving kidnapping, extortion, and death threats have been solved thanks to this type of expertise. For example, voice messages left on answering machines by kidnappers can be analysed in detail. This enables the identification of the criminal and their psychological state, as well as typical speech patterns, and even links to other incidents under investigation. Thus, speech evidence becomes a powerful tool in the hands of the judiciary, ensuring the inevitable punishment of criminals.

The Russian judicial practice shows that the results of forensic phonetic examination are effectively used in the investigation of economic crimes committed by a group of people. The members of a group use telephone communication to prepare, commit, and sell stolen property. In addition, the establishment and heading of a criminal community are mainly based on verbal communication. Let us take a practical example: when investigating a series of thefts of petroleum products

from several railway stations in the Chelyabinsk region, the investigators used the opinions of phonetic examinations as primary evidence. These tests identified the speakers based on the oral speech of a group of 35 suspects. In most cases, the audio recordings were the only evidence of the suspect's crime. There are several similar examples in the practice of Russian criminal proceedings.

There is hardly a criminal investigation into bribery or influence peddling that does not involve forensic phonetic analysis. Such criminal cases usually tend to have a high public profile and are followed with great interest by the general public.

In addition to the identification of a person by spoken language, Russian forensic phonetics deals with the issues of establishing the verbatim content of conversations with low-quality audio recordings (e.g., decoding "black boxes"), identifying the presence of audio editing or masking of a communicative event, establishing the fact of copying digital audio recordings, etc.

In Brazil, activities related to forensic phonetics cover a wide range of tasks and include identifying the speaker, analyzing the content of audio recordings, processing audio recordings to improve their intelligibility, analyzing sound signals associated with certain events (e.g., gunshots), linguistic analysis to determine the origin of the speaker, identifying signs of possible fraud in the processing of audio recordings, processing digital signals, examining communication systems, and analyzing systems for storing and compressing speech signals.

There are shining examples of high-profile criminal cases in which forensic phonetic analysis was the main evidence, including the above-mentioned case of the Minister of Labor, which was investigated in the early 1990s. The defense relied solely on denying that the suspect's voice on the recording belonged to Mr. Magri. The phonetic analysis became the main evidence of the Minister's guilt. Another example that received a wide media coverage was the disappearance of the mason Amarildo de Souza, involving several military police officers in Rio de Janeiro. The results of forensic phonetic examination confirmed the version that one of the police officers accused of the crime had spoken to another officer, posing as a drug dealer, to accuse him of kidnapping Amarildo. The third example is the case of the singer Marcelo Pires Vieira, known as Belo, who was sentenced in 2003 to 8 years in prison for drug trafficking and complicity in drug trafficking.

In India, there is a growing trend in cases where it is necessary to identify the speaker of recorded speech, such as in cases involving fake emergency calls, fraudulent phone transactions, obscene voicemails, and similar situations. According to the Global Scam Report 2021, India is the fourth most spammed country in the world. A survey conducted by Microsoft in 2021 found that India is considered a hub of scam call centers used for criminal purposes. Therefore, it is essential to create a database, especially in Indian English, to study speech characteristics. This will help in analyzing the intercepted and recorded data to narrow down the likely suspects based on speaker profiling.

2.3. Structure of Expert Institutions

In the Russian Federation, forensic phonetic investigation is carried out in various governmental organizations: the Ministry of Internal Affairs of the Russian Federation, the Ministry of Justice of the Russian Federation, and the Federal Security Service of the Russian Federation. They have a large group of experts, special methods of phonetic research, and specialised software. A specialist with a degree in engineering, programming, acoustics (an acoustician) or linguistics, speech therapy, philology (a linguist) can conduct forensic phonetic examinations. The forensic phonetic examination can be performed by one expert who has received the necessary qualifications in all types of analysis and has completed a certain number of internships and advanced training in the main expert institutions. After this, state experts receive a special document confirming the right to conduct phonetic examinations on a stand-alone basis, and then confirm their professional level every five years. A prerequisite for conducting forensic phonetic examinations is the availability of specialized equipment and phonetic software developed in Russia. It requires mandatory certification and regular metrological checks.

In Brazil, forensic experts can work at both the federal and state levels, and always on a competitive basis. The federal level is represented by the National Institute of Forensic Science (Instituto Nacional de Criminalística, INC)–a federal police agency responsible for the technical and scientific coordination of the entire federal forensic system. The Audiovisual and Electronic Forensic Science Service is accountable for analysing audio and video recordings, equipment, and electronic systems. This list also includes "speaker verification."

The capitals of the states, the federal districts and some cities in the country have a scientific and technical sector, which is a decentralized federal forensic unit technically linked to the National Institute of Forensic Science: the Technical Police Department (DPT) in the state of Bahia, the Renato Chaves Expertise Center in the state of Pará, the Forensic Science Institute (Instituto Geral de Perícias) in the state of Rio Grande do Sul. The federal forensic experts engaged in audio and video recording at the National Institute of Forensic Science, and decentralised units include specialists with degrees in electrical engineering, electronics, telecommunications, speech therapy, phonetics, etc. In addition, the Rio de Janeiro Public Prosecutor's Office has a DEDIT (Digital Evidence and Technology Division), which is part of the Security and Intelligence Coordination Center. Since 2009, the Brazilian Institute of Forensic Sciences (Instituto Brasileiro de Peritos, IBP) has been involved in major forensic investigations through forensic phonetic examination.

There are seven Central Forensic Science Laboratories (CFSL) in India, located in Chandigarh, Hyderabad, New Delhi, Kolkata, Pune, Bhopal, and Guwahati; only Chandigarh has a computerised speech laboratory and voice identification equipment. Some states have Forensic Laboratories, but only Tamil Nadu has a computerized voice laboratory with minimal equipment for voice identification.

2.4. Methodology and Strategy of Speaker Recognition

Speaker recognition by oral speech in the Russian Federation combines three types of research: perceptual auditory, acoustic, and linguistic analysis. This helps to make the forensic phonetic examination report more objective. The objectivity and reliability of the examination in the Russian Federation is increased by the fact that speaker recognition is carried out by two experts: a linguist and an acoustician. Let us consider the stages of identification research and the methods used at each stage. Speaker recognition based on audio recordings involves three main stages.

In primary perceptual auditory analysis, experts repeatedly listen to speech recordings and identify the speaker's individual features. They use the perceptual comparative analysis to compare the identified features in pairs. Introspection is used to identify the speakers' articulatory features. After the perceptive auditory analysis, experts make a preliminary decision on whether the criminal and the suspect are the same person.

The acoustic expert analyses the acoustic characteristics of the speech signal. These characteristics quantitatively reflect the physical parameters of speech signals, which are isolated and measured using computer-based mathematical software. One of the parameters usually analyzed by acoustic methods is the fundamental frequency of the speaker's voice. The expert also measures the duration of pronunciation and formants of vowels, which makes it possible to determine their articulatory characteristics. A similar analysis can be used to calculate the duration and articulation of sonorant consonants. The acoustician uses FFT spectral analysis (Fast Fourier Transform), formant structure analysis (for the vowels [a], [e], [i], [o], [u]), calculations of the time parameters of the duration of vowels and consonants, the fragmentary method, etc. The result of the acoustic analysis is based on more than 501 indicators of the physical parameters of speech in each 10-second segment of the disputed audio recording and the audio recording of the suspect's speech.

The linguist analyzes speech at three levels: the flow of speech, the level of words and phrases, as well as the phonetic level. Analyzing the flow of speech, the linguist describes voice timbre and pitch (or some abnormal qualities, such as hoarseness, dysphonia, hypophonia), speech rate, hesitation pauses, stuttering, etc. The linguist also identifies deviations from the speech norm, such as stuttering, rhotacism, a foreign accent, and regional dialects.

The expert analyzes the level of words and phrases using the lexical-semantic and semantic-syntactic methods. The first one evaluates individual word usage skills and preferences in choosing an inventory of lexemes, as well as the semantic coherence of lexical units. The second method helps to evaluate the speaker's individual preferences in selecting a syntactic construction demonstrating the organization of coherent speech in the communication process. Thus, the linguist makes a list of preferred vocabulary and syntactic constructions for a given speaker. For example, the speech of a petty thief typically contains obscene language and

simple, uncommon sentences, whereas the speech of a corrupt official often features professional vocabulary, expanded compound and complex sentences, and various types of syntactic connections between words.

Phonetic analysis primarily focuses on segmental features, including the use of vowels and consonants in coherent speech. The analysis also encompasses suprasegmental features of speech, including stress, intonation, and rhythmic word structure. The linguist listens carefully to the audio recording of speech several times and identifies individual articulatory and intonation features of the speaker's speech. For example, speech disorders such as rhotacism, lambdacism, sigmatism, dysarthria, etc. can be identified. Coarticulation and reduction of speech sounds and their relationship to speech norms are also analyzed. As a result of such an accurate analysis, the expert can identify the suspect, even when examining a very short speech recording.

All of the above stages refer to the research part of the forensic phonetic examination. The study covers separately the audio recording of the speech of the criminal and the suspect. Then comes the comparative part of forensic phonetic examination, which compares all identified perceptual, acoustic, and linguistic characteristics. During the comparative analysis, experts examine each distinctive feature of the speech of both the criminal and the suspect step by step.

In the final stage, the experts decide whether the suspect and the criminal are the same person. Sometimes the speech recording is so short and noisy that they cannot find enough individual characteristics to make a categorical conclusion. In this case, they say that they cannot decide whether the speech on the recording belongs to one person or another.

In Brazilian official forensic practice, forensic phonetic examinations are based on a series of research methods: perceptual auditory analysis and acoustic analysis with possible variations, for example, when a multidisciplinary team carries out the investigation: a speech therapist (perceptive auditory analysis) and an engineer (acoustic analysis, measurement of the physical parameters of speech signals), or when the investigation is carried out by a professional skilled in both areas of forensic phonetics. It is well known that the combination of the sensitive human ear with the visual and digital resources of acoustic analysis ensures the reliability of the decision taken.

According to Maria Lúcia de Castro Gomes & Andrea Alves Guimarães Dresch, the process of conducting forensic phonetic examination consists of the following stages. In the first stage, the quality of audio recordings is assessed, and they are deciphered, i.e., the verbatim content of conversation texts is determined. Experts discuss the need to establish the verbatim content of conversation texts with good speech intelligibility,[20] the possibility of re-recording and decoding a phonogram, and the associated responsibilities.

[20] Rehder, M. I. B. C., et al. (2014). Coincidências e divergências entre transcrição e textualização de áudios. *Revista CEFAC*, *16*(6), 1919–1927. (In Portuguese).

This is followed by the perceptual auditory analysis. Its results are used to transcribe oral speech using the International Phonetic Alphabet (IPA) symbols. Auditory analysis is not limited to describing the linguistic characteristics of speech sounds. It also identifies the characteristics of the speaker's voice, such as timbre, pitch, and strength. The Brazilian Portuguese vocal profile analysis scheme is used to record the results of this analysis (*see* Figure 1). This standard allows for the detailed recording of the speaker's oral speech characteristics, i.e., the deviations from the average "normal" values.

QUALIDADE VOCAL	**PRIMEIRA PASSADA**		**SEGUNDA PASSADA**							
	Neutro	**Não neutro**	**AJUSTE**	**Moderado**			**Extremo**			
				1	**2**	**3**	**4**	**5**	**6**	
A. ELEMENTOS DO TRATO VOCAL										
1.Lábios			Arredondados/protraídos							
			Estirados							
			Labiodentalização							
			Extensão diminuída							
			Extensão aumentada							
2. Mandíbula			Fechada							
			Aberta							
			Protraída							
			Extensão diminuída							
			Extensão aumentada							
3.Língua ponta/lâmina			Avançada							
			Recuada							
4. Corpo de língua			Avançado							
			Recuado							
			Elevado							
			Abaixado							
			Extensão diminuída							
			Extensão aumentada							
5.Faringe			Constrição							
			Expansão							
6.Velofaringe			Escape nasal audível							
			Nasal							
			Denasal							
7. Altura de laringe			Elevada							
			Abaixada							
B. TENSÃO MUSCULAR GERAL										
8. Tensão do trato vocal			Hiperfunção							
			Hipofunção							
9. Tensão laríngea			Hiperfunção							
			Hipofunção							
C. ELEMENTOS FONATÓRIOS										
	AJUSTE		**Presente**			**Graus de escala**				
			Neutro	**Não Neutro**		**Moderado**			**Extremo**	
						1	**2**	**3**	**4**	**5**	**6**
10. Modo de fonação	Modal										
	Falsete										
	Crepitância/ *vocal fry*										
	Voz crepitante										
11. Fricção laríngea	Escape de ar										
	Voz soprosa										
12.Irregularidade laríngea	Voz áspera										
DINÂMICA VOCAL		**Neutro**	**AJUSTE**			**Moderado**			**Extremo**		
						1	**2**	**3**	**4**	**5**	**6**

D. ELEMENTOS PROSÓDICOS									
13.*Pitch (f0)*	Habitual		Elevado						
			Abaixado						
	Extensão		Diminuída						
			Aumentada						
	Variabilidade		Diminuída						
			Aumentada						
14.*Loudness (intensidade)*	Habitual		Aumentado						
			Diminuído						
	Extensão		Diminuída						
			Aumentada						
	Variabilidade		Diminuída						
			Aumentada						
15. Tempo									
Continuidade			Interrompida						
Taxa de elocução			Rápida						
			Lenta						
16.OUTROS ELEMENTOS									
Suporte respiratório			Adequado						
			Inadequado						

Ocorrências em curto termo ( )quebras ( ) instabilidades ( ) diplofonia ( ) tremor
Para ajustes de ocorrência intermitente assinalar (i)

Figure 1: **Vocal profile protocol of Brazilian Portuguese**

Acoustic analysis is conducted using such standard phonetic analysis programs as Praat, Cool Edit, Sound Forge, etc. These programs allow the measurement of the main physical parameters of the speech signal, and the compilation of the formant structure of vowels.

The results of the conducted research are formalized in an official report presented in court.

Notably, the methods used in the identification research of oral analysis in Russia and Brazil are quite similar. Nevertheless, a detailed study of the methods and techniques of conducting forensic phonetic examination in both countries shows that they can be mutually enriched by introducing some research methods adopted in another country. Thus, it seems promising to formulate a vocal profile protocol similar to the one above for the Russian language and to use it in phonetic examinations in Russia.

The creation of a substantial database of audio recordings of Indian oral speech is considered by Indian scholars and experts working in the field of oral identification to be a prerequisite for further research in this area. The complexity of the language

and the size of the Indian population present numerous problems in speech analysis, making it crucial to conduct research at various sub-levels. The creation of an Indian speech database would facilitate the study of speaker characteristics, thereby enabling the evaluation of the feasibility of using speaker-specific variations to differentiate individuals within a large population of speakers. This database could also quantify the articulatory-acoustic features of speakers and aid in examining whether diachronic changes can be a factor of speaker idiosyncrasy. In fact, the automation methods of speaker identification in Russian expert institutions are based on the processing of a range of data on the voices and speech of the population.

2.5. Current Issues of Forensic Phonetic Examinations

In Russian expert practice, the recognition of a speaker using synthesized speech or voice-changing devices is most urgent. Cases of fraud committed using synthesized oral speech of a person known to the victim (a close relative, work colleague, manager, etc.) have become more frequent. Sophisticated speech technologies sometimes serve as a tool to conceal crimes, and experts need new methods and solutions to combat fraudsters. Synthesized speech created in a text-to-speech conversion system can be considered as a written text. Therefore, a forensic expert can use a method to determine authorship based on written speech. If a criminal uses a voice-altering application, a professional expert can detect the method of disguise and perform reverse acoustic transformations to return the speech signal to its original state, allowing for further recognition. However, sometimes such transformations are not possible.

The qualification of a forensic phonetic expert is the most discussed issue in the scientific literature. The interdisciplinary nature of forensic phonetic examination provides a high level of expert knowledge in several disciplines, including forensic expertology, criminalistics, linguistics, phonetics, and acoustics. As Elena Galyashina notes,

> Forensic linguistics has expanded in litigation practice, but the problem of finding correlation between the specializations of linguists and criminalists (forensic experts) persists.

The results of forensic phonetic examination should comply with the principle of scientific objectivity and verifiability of results. The procedure for applying each method should be described in detail. Thus, experts and scientists agree on the need for a uniform state standard for determining the competence of experts.

For Brazilian expert practice, the issue of forensic phonetics terminology is controversial. Official forensic investigation bodies in Brazil use the nomenclature of "speaker verification" (in the practice of forensic phonetic examination in the Russian Federation, the term "verification" means the comparison of unknown speakers in two or more disputed audio recordings. Further in the text, the authors of this paper use the terms speaker recognition and verification as synonyms).

> The nature of human speech and the existing analysis procedures do not allow for the individualization of a speaker from a large database with the precision that can be achieved using molecular genetics and fingerprinting. Therefore, in Brazilian forensic terminology "identification" means that the characteristics presented in the evidence are found only in the author with a certain degree of precision and taking into account the population of the planet.[21]

The qualification and specialization of a forensic phonetics expert is also relevant in Brazil. According to available data, the Federal Police employs mainly engineers, the Rio de Janeiro Public Prosecutor's Office employs only speech therapists, and the Institute of Forensic Science of the State of Paraná employs a somewhat more multidisciplinary groupwhich includes engineers, speech therapists, and IT specialists.

Specialists in forensic phonetics in Brazil are working on the development of the following topics: voice identification of siblings,[22] phonation disorders as an identification feature,[23] the role of sociophonetics in speaker recognition,[24] and dialectal features in speech as the speaker's identification feature.

In Russia and Brazil, in particular, the development of forensic linguistics in the broad sense began with forensic phonetics. The first academic papers published in the first two issues of the Brazilian scientific journal Language and Law / Linguagem e Direito deal with forensic phonetics. A similar global trend can be observed in the development of the theory of forensic linguistics. We believe that the reason for this is that the methods of studying objects in forensic phonetics are most closely aligned with the techniques used in traditional forensic investigations.

An analysis of the role of forensic phonetics in criminal proceedings in Russia and Brazil reveals several similarities. Forensic phonetics is certainly an interdisciplinary field that requires both highly qualified specialists and specialised equipment and software. As practice and accepted legal recommendations demonstrate, both countries hold the position that no special knowledge in forensic phonetics is required to determine the verbatim content of the text of phonograms containing intelligible speech. Both countries have a binary structure of the expert opinion"he/she," which does not imply any gradation options.

---

[21] Gomes, M. L. C., Richert, L. C., & Malakoski, J. (2012). Identificação de locutor na área forense: A importância da pesquisa interdisciplinary. *Encontro do CELSU*, *24*(6), 1–13. (In Portuguese).

[22] San Segundo, E., & Gómez-Vilda, P. (2014). Evaluating the forensic importance of glottal source features through the voice analysis of twins and non-twin siblings. *Language and Law/Linguagem e Direito*, *1*(2), 22–41.

[23] Gómez, P., San Segundo, E., Mazaira, L. M., Álvarez, A., & Rodellar, V. (2017). Using dysphonic voice to characterize speaker's biometry. *Language and Law/Linguagem e Direito*, *1*(2), 2–66. (In Portuguese).

[24] Gonçalves, 2021.

At the same time, there are differences in the experience of conducting phonetic examinations. The exchange of expertise enriches phonetic analysis methods in both countries. For example, in Russia, the speaker's oral speech is not transcribed using IPA symbols, although other countires' phoneticians indicate the need for this. In Russian practice, a significant portion of forensic phonetic examination is devoted to analysing the communicative, lexical, and syntactic features of speech. In Brazilian practice, the specified features of speakers are not analysed, which is explained by the composition of the experts. In Brazil, this type of examination is more often carried out by speech therapists and engineers than by linguists.

3. Authorship Attribution

Authorship attribution of texts has become particularly important because of the vast number of written works produced by Homo Scribens. Never before has humanity exchanged so many texts. The intensification of Internet communication has led to a number of security issues: a virtual interlocutor on the Internet can pretend to be another person, in order to cause material or moral harm to the interlocutor; the volume of dangerous content, such as cyberbullying, incitement to hatred or hostility, is increasing.[25] Moreover, when investigating a criminal case, it becomes necessary to identify the author of the text in cases of blackmail, kidnapping for ransom, anonymous letters containing a threat, texts inciting to depraved actions, suicide, letters from suicides (when staged or driven to suicide), etc.

An analysis of the English scientific literature shows that the number of works on authorship attribution based on linguistic analysis has increased rapidly in the last five years. Over the last three years, the number of forensic authorship examinations in Russia has increased by 20%.[26] The relevance of authorship examination is also growing because most texts in circulation are digital, virtual, and not subject to recognized handwriting/graphometric examination.

Sapir writes:

> Each of us has his own individual style [...]. There is always an individual method employed in combining words into groups, and these groups into larger units.[27]

This idea perfectly summarizes the basic premise of authorship attribution that different people use language to some extent differently. This unique set of ways of using language is called an idiolect or individual style (as opposed to social styles or dialects).

---

[25] Panicheva, P., & Litvinova, T. (2019). Authorship attribution in Russian in real-world forensics scenario. In *Proceedings of the International Conference on Statistical Language and Speech Processing* (pp. 299–310). Springer.

[26] Khairullova, E. T. (2022). Autoratic research: Practice and prospects of application. *Law and Legislation*, *7*, 221–223. (In Russian).

[27] Sapir, E. (1927). Speech as a personality trait. *American Journal of Sociology*, *32*(6), 892–905.

3.1. Categories of Cases and Judicial Practice

The term "authorship" was introduced into Russian legal practice in the 1970s by S. V. Vul. Authorship attribution of a text was initially mentioned in copyright protection proceedings, identifying signs of plagiarism, in cases of protecting honour, dignity and business reputation, as well as slander. As for criminal proceedings, authorship should be most often attributed in cases of extremism, threats, incitement to suicide, murder, etc. Authorship attribution of a suicide note has become necessary evidence, confirming the version of a staged suicide. For example, during the investigation, it became necessary to establish the factual circumstances of whether female citizen A. was actually the author of the suicide note. The authorship examination revealed specific diagnostic features of the author of the note, namely, that the author was a male and not a native Russian speaker. In fact, this helped to identify the suspect, who later confessed to the murder and staged the suicide.

The proliferation of anonymous extremist content on the Internet presents a challenge to law enforcement officers in authorship attribution of these texts. To this end, authorship expertise has become highly sought after in the investigation of criminal cases of an extremist nature. For example, during the investigation it was discovered that extremist materials were published from the IP address of a work computer accessed by more than 30 people. In this case, a diagnostic authorship study, along with determining the demographic characteristics of the author (gender, age, ethnicity, territorial affiliation, and level of education), allowed for a significantly narrowed circle of suspects. The subsequent authorship examination identified the criminal.

The following case illustrates the use of forensic authorship expertise to prove a criminal case in Brazil. Nicea Pitta, the First Lady of São Paulo, accused the businessman Jorge Yunes of writing and sending her a death threat letter. The death threat letter received by Nicea was dated March, 3, 2000 and consisted of one page filled with a total of sixty-six words cut out of a magazine or newspaper and interspersed with handwritten words. Experts had to determine whether Jorge Yunes was indeed the author of the letter. Recordings of interviews with Jorge Yunes about the accusation were provided as comparison material. Linguists carried out an attribution study, which concluded that Jorge Yunes was not the author of the anonymous letter. It is noteworthy that the analysis was based on a new field in authorship studies, discourse analysis, which favours the study of the fundamental, deepest, and most abstract level of the organization of meaning. This type of analysis was chosen because the experts compared texts of different genres and because the interview was prompted by a letter, which in itself created a coincidence of topics. The outcome of the case was a counterclaim by Jorge Yunes for defamation, which was satisfied, and Nisea Pitta was sentenced to six months in prison, which was replaced by a fine.

In the context of the Indian judiciary, there is no reference to the use of text-based identification techniques in criminal investigations. However, given the predominant

shift in communication to messenger platforms, it is evident that this area requires urgent investment and development.

3.2. Methodology and Strategy of Authorship Attribution

There are two applicable main approaches to solving the problem of authorship attribution: stylistic (qualitative, idiographic, i.e., based on the analysis of individual features of the text chosen subjectively by the researcher for each attribution situation) and stylometric (quantitative, based on the analysis of many stylometric features of the text using statistical methods and/or machine learning algorithms). According to the accurate observation of the English researcher D. Wright, these approaches are at best developing in parallel, and at worst, competing with each other.[28]

In a broad sense, forensic authorship is an interdisciplinary field of knowledge that studies the theoretical and practical aspects of authorship attribution of products of human activity and artificial intelligence.[29] Authorship examination addresses the issue of identifying a person based on a written text, including the author's individual features and the conditions under which the work was created.

The tasks solved by an expert can be divided into two large groups: 1) identification tasks related to the authorship attribution of a text by conducting the separate and comparative analysis of the features that appear in the disputed text and texts–samples of the suspect's written speech; 2) diagnostic tasks related to determining the age, gender, individual and personal features, the level of communicative competence, speech culture, and the sphere of professional activity of the author of the text.[30] To solve both types of tasks, Russian expert practice applies specially developed, tested, and approved methods.[31,32] It can be said without prejudice that the results of expert authorship studies are recognized as a weighty argument in Russian judicial practice.

In Russia, the work of an authorship expert is based on theoretical and experimental research in criminology and forensic examination, forensic authorship studies, linguistic statistics, physiology, biomechanics, psycholinguistics, psychology, cybernetics, and extensive factual material based on the study and generalization

---

[28] Wright, D. (2014). *Stylistics versus statistics: A corpus linguistic approach to combining techniques in forensic authorship analysis using Enron emails* (PhD thesis). University of Leeds.

[29] Sokolova, T. P. (2019). Forensic author identification as an interdisciplinary field of knowledge. *Courier of the Kutafin Moscow State Law University*, *5*(57), 132–142. (In Russian).

[30] Nazarova, T. V., & Gromova, A. V. (2016). Objects and tasks of linguistic and authorship examinations carried out in forensic units of the internal affairs agencies of the Russian Federation. *Forensic Examination of Belarus*, *1*, 43–46. (In Russian).

[31] Nazarova & Gromova, 2016.

[32] Izotova T. M., Plotnikova A. M., Kuznetsov V. O., & Kryuk E. K. (2020). *Guidelines for conducting forensic authorship examinations* (pp. 207–210). Publishing House of the RFCFS of the Russian Ministry of Justice. (In Russian).

of expert practice. The methodological approach is based on the results of special experiments using mathematical modeling, which determines their reliability and scientific validation.[33] A specific research methodology can be hardly shown in authorship studies; the expert selects the necessary research methods based on the provided object. Therefore, we will only list linguistic analysis methods used in identifying and diagnosing the author of the text: the method of analysing the structural organisation of the text, the method of lexical-semantic analysis, semantic-syntactic analysis, component analysis, content analysis, functional-stylistic analysis, and quantitative and corpus linguistic methods.

Russian authorship studies widely use computer methods in solving the problem of text authorship. This context frames the concepts of "stylemetry," "identification idiolectology,"[34] as well as the RusIdiolect resource, unique for corpus linguistics, intended for experimental studies of the idiolect of an "ordinary" native speaker. The "computational" approach to identifying the author of a text generally consists of constructing a classifier, whose input is the numerical values of various automatically extracted quantifiable parameters of a text (unigrams and n-grams, i.e., sequences of n-elements–symbols (letters, punctuation marks, numbers, etc.), words, parts of speech, etc.); less often–syntactic and semantic parameters; and whose output is the class of the object (text), i.e., its attribution to a particular author.[35] Although objective, quantitative and corpus linguistic methods apply mainly to large texts (more than 2,500 words), in practice, however, such texts rarely become evidence in criminal investigations. Therefore, these methods are rather an additional tool that extends the capabilities of the authorship researcher.

In Brazil, much of what is being done in forensic authorship is carried out *ad hoc*, i.e., in each specific case, and not as per a standard or protocol. Dayane Celestino de Almeida, author of a thesis on styles and identities in the context of authorship attribution in the forensic field,[36] writes that many attribution methods have been developed to attribute the authorship of texts in the historical, religious, and literary sphere, ignoring the specific features of forensic contexts. In the context of stylometric studies, Brazilian linguists analyze the following linguistic markers: word or sentence size, n-gram and "collocation" frequency, lexical or grammatical word frequency, lexical "density" or "richness," upper and lower case usage, morphosyntactic features, abbreviation styles, punctuation, greeting and closing styles. Repetition is needed to use the specified markers; therefore, stylometrics is applicable to long texts consisting of thousands of words. At the same time, disputed texts are usually

33 Galyashina, 2018.

34 *Id*., p. 29.

35 Litvinova, T. A., & Gromova, A. V. (2020). Computer technologies in forensic authorship examination: Problems and prospects of use. *Bulletin of Volgograd State University*, *19*(1), 77–88. (In Russian).

36 Almeida, 2015.

short in criminal proceedings and the linguistic material is scarce. Furthermore, the choice of a certain marker and its validity is not always clear.

The expert practice of Brazilian authorship studies also uses the concept of "idiolect," which is based not only on the repeated use of a particular linguistic variable, but rather on their combination–"constellation of variables."

Brazilian linguists also make use of Forensic Discourse Analysis, which is new to the global practice of authorship studies. Its main task is to identify uses of language that are "out of place." For example, by analyzing a written record of an interrogation, you can get information that it is a forgery if there are traces of police discourse, clichés in what should be taken literally from the speech of an "ordinary citizen."

Practitioners point out that it is better to think of quantitative and qualitative analysis as stages of analysis rather than different methods of analysis. It seems reasonable to assume that the different lines of analysis and the two main methodological approaches–quantitative and qualitative–have their own strengths and weaknesses and can be used as complementary ones.[37]

They also point out that cases of using authorship expertise in the investigation of real criminal cases are extremely rare because of the lack of research and publications in this area resulting in the lack of knowledge among law enforcement officers about what a linguist can or cannot do in relation to a text ofquestionable or unknown authorship.[38]

3.3. Current Issues of Authorship Examination

The challenges of authorship verification in Russia include the issues related to the quantity and quality of forensically significant texts. Most texts subject to authorship examination contain less than 1,000 words. Existing expert methods in Russia allow working with texts of 200 words or more, but in such cases a high degree of comparability of the analyzed text with comparative samples is crucial. A forensically significant text may be written by a group of authors. There are various ways in which two or more people can be involved in the production of a text: fragments of the text belong to different authors; the text is written by one but edited by another author; two (or more) authors took an equal part in the production of the text, etc. In addition, the features of the author's text can be imitated or disguised. The author of the text may imitate a sociolect they believe to be typical of persons of a certain group (gender, age, social class), to which they do not belong, or an idiolect of a specific person. There are also cases of idiolect masking which does not focus on any standard to hide one's identity (for example, by reducing the level of literacy). This also includes the problem of synthesised written speech (ChatGPT, YaGPT 2). The "Write it in another way" function allows for significant changes to the text style while preserving the content.

---

[37] Almeida, 2015, p. 5.

[38] *Id*., p. 5.

Theoretical research conducted by Russian linguists in the field of authorship examination study a new type of text. This is so-called keyboard-mediated speech results from the transition of a person's oral speech to written speech. To put it simply, the author writes down his/her oral speech in written characters. These are mainly unprepared, spontaneous, and unedited written statements, closely structured in terms of lexical and grammatical means to oral speech. The speech rate of Internet communication texts is close to the oral variety of speech. Various non-verbal symbols, signs (emoticons), abbreviations, and multiple reductions are actively used to express intonation, helping the communicator to express emotional coloring. Moreover, the creation of Internet communication texts using a keyboard and writing a regular written text are completely different psychological and physiological processes.[39]

Brazilian scholars believe that not all authorship cases should be solved absolutely the same way, using the same linguistic markers or idiosyncrasies. Preliminary research is needed to determine whether a linguodiscursive element is valid and reliable for identifying the authorship of the text. The validity and reliability of the method can only be determined based on conducted and, ideally, published studies. The work of linguists as experts and consultants will have more weight if verifiable and reproducible methods support it. There is no single model that would be superior to others between quantitative and qualitative research. However, there is a tendency to use hybrid models in which "quali" and "quanti" are analysis stages rather than different paradigms.

In Portuguese-speaking countries, academic research on forensic authorship attribution is still in its early stages, despite recent growth. This does not mean that no professionals are working on these issues in practice. However, they are still few and far between. and there is a lack of: a) academic research in Portuguese that serves as a basis and support for their work; or b) reports on the methods they use in real cases they have dealt with.

The analysis of forensic authorship studies in the countries under study allows us to conclude that this type of expertise is becoming increasingly relevant. The growing demand for this type of expert research is explained by the large volume of electronic texts circulating on the Internet. Taking into account the similar structure of the Russian and Portuguese languages (inflectional) and the similar scientific approaches used by authorship researchers to identify a person through a written text, we believe the exchange of analysis methods, as well as linguistic research collaborations in authorship studies, contribute to identifying the most effective approaches to solving the existing problems.

---

[39] Galyashina, E. I., & Privodnova, E. V. (2006). Authorship examination in Russian legal proceedings. *Lex russica*, *65*(4), 755–761. (In Russian).

Conclusion

Forensic linguistics is a rapidly developing field in many countries around the world. Worldwide legal practice demonstrates that the use of linguistic methods for identifying individuals through oral and written speech is an effective tool in the hands of law enforcement agencies for investigating and detecting criminal cases. For example, in cases involving bribery, extortion, threat, and kidnapping, an audio recording or text is often the only evidence. In such cases, the examination of a linguistic expert becomes crucial for identifying the person in question. A comparative analysis of the use of linguistic expertise in Russia, Brazil, and India has demonstrated the significant demand for such knowledge. In Russia and Brazil, expert examinations to identify individuals through their oral and written speech are a common practice and a methodologically sound approach. These countries have established special institutions where such expert examinations are conducted and training in these specialised fields is facilitated. It has been observed that there exists a considerable potential for the exchange of methodological advances between the two nations. Indian jurists have repeatedly emphasised the need to incorporate linguistic expertise into the legal domain. This is particularly important given India's vast linguistic and dialectal diversity. However, it is regrettable that the use of forensic linguistics within the Indian judicial system remains limited at present. In this regard, it is considered beneficial to draw on the experience of Russia, Brazil and other countries in establishing forensic linguistics in India. This should include the creation of educational programmes in forensic linguistics, the conduct of scientific research in the field of personality identification through oral and written speech, and other relevant areas. This is particularly important given the potential of linguistic knowledge to contribute to the effectiveness of the administration of justice.

References

Coulthard, M., Johnson, A., & Wright, D. (2016). *An introduction to forensic linguistics: Language in evidence*. Routledge. https://doi.org/10.4324/9781315630311

Figueiredo, D. (2014). Discurso, gênero e violência: uma análise de representações públicas do crime de estupro. *Language and Law/Linguagem e Direito*, *1*(1), 141–158. (In Portuguese).

Galyashina, E. I. (2018). Forensic linguistics in Russia: State of the art and new challenges. *Theory and Practice of Forensic Examination*, *13*(4), 28–37. (In Russian).

Galyashina, E. I., & Privodnova, E. V. (2006). Authorship examination in Russian legal proceedings. *Lex russica*, *65*(4), 755–761. (In Russian).

Gibbons, J., & Turell, M. T. (Eds.) (2008). *Dimensions of forensic linguistics*. Equinox Publishing. https://doi.org/10.1075/aals.5.19con

Golev, N. D. (1999). The legal aspect of language in linguistic coverage. *Legal Linguistics*, *1*(1), 12–59. (In Russian).

Gomes, M. L. C., Richert, L. C., & Malakoski, J. (2012). Identificação de locutor na área forense: A importância da pesquisa interdisciplinary. *Encontro do CELSU*, *24*(6), 1–13. (In Portuguese).

Gómez, P., San Segundo, E., Mazaira, L. M., Álvarez, A., & Rodellar, V. (2017). Using dysphonic voice to characterize speaker's biometry. *Language and Law/Linguagem e Direito*, *1*(2), 2–66. (In Portuguese).https://ojs.letras.up.pt/index.php/LLLD/article/view/2431

Gonçalves, C. S., & Brescancini, C. R. (2017). Considerações sobre o papel da sociofonética na comparação forense de locutores. *Language and Law/Linguagem e Direito*, *1*(2), 67–87. (In Portuguese).

Gonçalves, J. S. S. (2021). Práticas de Análise em Linguística Forense. *Language and Law/Linguagem e Direito*, *8*(1), 120–124. (In Portuguese).

Hollien, H. (1990). *The acoustics of crime: The new science of forensic phonetics*. Springer.

Izotova T. M., Plotnikova A. M., Kuznetsov V. O., & Kryuk E. K. (2020). *Guidelines for conducting forensic authorship examinations*. Publishing House of the RFCFS of the Russian Ministry of Justice. https://doi.org/10.30764/978-5-91133-219-8-2020-12. (In Russian).

Khairullova, E. T. (2022). Autoratic research: Practice and prospects of application. *Law and Legislation*, *7*, 221–223. https://doi.org/10.24412/2073-3313-2022-7-221-223. (In Russian).

Litvinova, T. A., & Gromova, A. V. (2020). Computer technologies in forensic authorship examination: Problems and prospects of use. *Bulletin of Volgograd State University*, *19*(1), 77–88. https://doi.org/10.15688/jvolsu2.2020.1.7. (In Russian).

Nazarova, T. V., & Gromova, A. V. (2016). Objects and tasks of linguistic and authorship examinations carried out in forensic units of the internal affairs agencies of the Russian Federation. *Forensic Examination of Belarus*, *1*, 43–46. (In Russian).

Panicheva, P., & Litvinova, T. (2019). Authorship attribution in Russian in real-world forensics scenario. In *Proceedings of the International Conference on Statistical Language and Speech Processing* (pp. 299–310). Springer.

Rehder, M. I. B. C., et al. (2014). Coincidências e divergências entre transcrição e textualização de áudios. *Revista CEFAC*, *16*(6), 1919–1927. (In Portuguese).

Rose, P. (2002). *Forensic speaker identification*. CRC Press. https://doi.org/10.1201/9780203166369

Rusman, G., D'Orio, E., Popova, E., & Kipouras, P. (2023). Features of the application of digital technology in criminal proceedings of the BRICS countries. *BRICS Law Journal*, *10*(1), 35–58. https://doi.org/10.21684/2412-2343-2023-10-1-35-58

San Segundo, E., & Gómez-Vilda, P. (2014). Evaluating the forensic importance of glottal source features through the voice analysis of twins and non-twin siblings. *Language and Law/Linguagem e Direito*, *1*(2), 22–41.

Sapir, E. (1927). Speech as a personality trait. *American Journal of Sociology*, *32*(6), 892–905.

Sokolova, T. P. (2019). Forensic author identification as an interdisciplinary field of knowledge. *Courier of the Kutafin Moscow State Law University*, *5*(57), 132–142. https://doi.org/10.17803/2311-5998-2019-57-5-132-142. (In Russian).

Toppo, R., & Sinha, S. (2023). Evaluating and accessing the scope of forensic linguistics in a multilingual context in India. *Language in India*, *23*(8), 113–159.

Information about the authors

Julia Morozova (Chelyabinsk, Russian Federation) – Associate Professor, Department of Criminal Procedure, Criminalistics and Forensic Examination, South Ural State University (National Research University) (76 Lenina Ave., Chelyabinsk, 454080, Russian Federation; e-mail: kudriavtcevaya@susu.ru).

Natalie Papoyan (Chelyabinsk, Russian Federation) – Associate Professor, Department of Linguistics and Translation Studies, Department of Criminal Procedure, Criminalistics and Forensic Examination, South Ural State University (National Research University) (76 Lenina Ave., Chelyabinsk, 454080, Russian Federation; e-mail: papoiannp@susu.ru) – **corresponding author**.

Niteesh Kumar Upadhyay (Gandhinagar, India) – Associate Professor, School of Law, Dhirubhai Ambani University (DA-IICT Campus, Near Reliance Cross Rd., Gandhinagar, Gujarat, 382007, India) and Research Fellow, Tashkent State Law University, Uzbekistan (e-mail: niteeshkumar_upadhyay@dau.ac.in).