

EXPLORING LEGAL ENGLISH: WAYS TO INTEGRATE AI IN YOUR RESEARCH

Olga B. Bagrintseva
Olga M. Litvishko¹
Elena V. Ryumkova

Received 11.02.2026.

Revised 12.04.2026.

Accepted 01.06.2026.

Keywords:

Enterprise, Competition, Market, Market segment, Low costs, Competitive strategy, Competitive advantage, Ways to compete.

Original research



ABSTRACT

The research examines the current possibilities and prospects of using artificial intelligence (AI) in the linguistic analysis of legal English, which is the international language of law. Particular attention is paid to the genre and terminological heterogeneity of legal texts, which places high demands on the adaptation of AI tools for accurate and contextually correct analysis. The research covers an overview of existing technologies, including automatic recognition and classification of terms, machine translation, automatic annotation, thematic and stylistic analysis, and generative models for educational purposes. The descriptions of the Sonix and ContractCrab platforms are of practical importance, demonstrating different approaches to working with audio and text legal data, which expand the horizons of collecting and processing linguistic information. The authors highlight the benefits of automation, improving the accuracy and speed of research, and draw attention to the limitations associated with language specifics and ethical aspects of using AI. The research emphasizes the need to integrate AI, considering the specifics of legal discourse and the prospect of creating intelligent support systems for lawyers, linguists, and participants in international communication.

© 2026 SPECTRUM Journal of Social Sciences

1. INTRODUCTION

Active and widespread implementation of Artificial Intelligence (AI) in all areas of life transforms the very essence of communication and cognition, offering new mechanisms for information processing and decision-making. AI technologies are changing traditional work models in education, medicine, business, and law, opening up opportunities for automating routine processes and stimulating innovation. Linguistics is no exception. The influence of AI technologies on conducting linguistic research into the language of law and legal communication is becoming increasingly noticeable. Being the international language of law and the language of international law and international legal interaction, English requires the integration of AI tools to ensure the accurate analysis of specialized texts.

There is no denying that legal English is characterized by significant diversity in genres and textual and linguistic features; it is clearly distinguished among other professional languages by a special linguistic identity, which underscores the necessity of adopting new methods of research. Traditional methods of analysis often prove to be insufficiently effective for processing large amounts of textual information. Implementation of AI technologies opens up new horizons for the research and practical application of the obtained results, allowing for a significant increase in the accuracy and speed of analysis. This research aims to explore the possibilities and ways of using AI in the linguistic analysis of legal English. Particularly, the research examines how AI can improve the research into understanding and interpretation of legal documents, as well as what new methods of text analysis can be developed using it. The

¹ Corresponding author: Olga M. Litvishko
Email: olitvishko@yandex.ru

influence of advanced AI technologies on the development of linguistics is apparent, especially in terms of natural language processing and the creation of intelligent analysis systems that are also used in the legal field, facilitating work with legal information. However, legal English places special demands on AI tools: such features of legal English as high formality, terminological complexity, and unique syntactic constructions (Litvishko et al., 2025b) require the adaptation of models for accurate interpretation and processing of texts. The practical significance of this research lies in the possibility of improving translation, education, and law enforcement through the use of AI, which will positively affect professional linguistics, legal practice, and intercultural communication at the international level.

2. METHODOLOGY

The research aims to identify the specifics of the use of AI in studying legal English and assessing the prospects for the development of technologies in this area. To achieve this goal, the following objectives are set:

- To study the existing AI technologies used in linguistic research;
- To analyze the features of AI tools applicable to linguistic analysis;
- To evaluate the advantages and limitations;
- To outline the prospects and limitations of AI integration in the research of legal English.

To achieve the goals set, the authors applied analysis, synthesis, generalization, corpus studies, and the method of quantitative and computer linguistics. The initial analysis of the theoretical material made it possible to clearly structure the information and identify the general parameters under study. The use of analysis and synthesis methods contributed to highlighting the key ideas in the studied scholarly publications, determining general and specific findings in the analyzed works. Elements of corpus and quantitative and computer linguistics are used to demonstrate the ways AI is employed in studying legal English. Generalization of the research results led to formulating the integral conclusions and revealing the limitations of AI integration in the research of legal discourse.

The research is based on the theoretical review of scholarly publications on relevant issues, functional benchmarking of AI tools, and linguistic analysis of various legal texts.

3. RESULTS

In practice, AI has already been successfully integrated into the analysis based on processing legal English texts to solve various research problems (Leonardi & Giampieri, 2025; Litvishko et al., 2025a). For instance, automatic recognition and classification systems of legal terms facilitate the development of specialized lexicons

and databases, increasing the accuracy and speed of searching for relevant information. Neural network models make it possible to effectively identify legal terminology and ambiguous words in the texts of court decisions, contracts, and regulations, which greatly simplifies linguistic analysis and translation (Armour et al., 2020). Another example is the use of machine translation and automatic text annotation technologies, which makes legal information accessible to professionals and a wide audience who do not speak English perfectly. Additionally, AI tools are used to identify patterns of argumentation and the structure of judicial precedents, which contribute to the development of intelligent decision support systems and educational platforms in the legal field (Tu et al., 2024). Finally, AI generative models are used in designing educational materials and practical cases for learning legal English (Storey et al., 2025). They can adapt complex text to students' language proficiency, automatically generate exercises, and create interactive dialogues to master professional terminology and legal communication skills.

Let us briefly describe some platforms and services that may be used to facilitate the linguistic research on legal language. Most AI tools listed below are the instruments used by professional lawyers and law students in their routine professional activities or legal research. Some of the tools are used by lay people to get free legal advice and assistance, or have a complex legal text simplified, or a legal document written. The authors chose some services among a great diversity of available AI tools (the website "Trending AI Tools (n.d.)" offers 174 results in the category Legal and Compliance) to show that their use is not limited solely to the field of law but may also be integrated into linguistic research. Below is the list of AI tools with their features related to linguistic analysis:

1. AI Lawyer (<https://ailawyer.pro/>) is used to get answers to legal questions, create and compare documents, formulate legal issues, and simplify complex concepts for a lay audience;
2. CaseText (<https://casetext.com/>) is applied when conducting an in-depth analysis of judicial practice and identifying linguistic patterns in relevant arguments;
3. ContractCrab (<https://contractcrab.com/>) is used to reduce the volume of legal documents while preserving key information. This tool helps analyze the structure and prevailing language in legal texts;
4. DeepL (<https://www.deepl.com/>) provides accurate machine translation of legal texts while preserving terminological specifics;
5. DoNotPay (<https://donotpay.com/>) can provide automated legal advice and assistance in drafting documents. It contains extensive material on the practical aspects of using legal English in dialogues and statements;
6. Eigen Technologies (<https://www.eigentech.com/>) is used to analyze lexical and semantic features of legal texts, structure and classify legal terminology, and study stylistics and dynamics of legal documents

- through comparison and automatic summarization tools;
7. Ghostwriter Legal (<https://ghostwriter-ai.com/>) is used for automating the style and tone of professional legal texts, exploring stylistic parameters, and maintaining language uniformity;
 8. Ivo (<https://www.ivo.ai/>) helps to study the structure of legal texts and communication models in jurisprudence;
 9. Law.co (<https://law.co/>) can interpret, compress, and analyze contracts and help in studying terminological features in legal contracts;
 10. Lawgeex (<https://www.lawgeex.com/>) is used to automate contract verification and identification of legal risks, which makes it possible to study the structure of contractual vocabulary and typical formulations;
 11. LawGeex (<https://www.lawgeex.com/>) helps analyze the accuracy and relevance of legal formulations;
 12. Legal Robot (<https://legalrobot.com/>) is used to conduct automated analysis of documents that translates complex legal formulations into simple language and helps identify risks and key conditions. It also helps analyze terminology and simplify legal texts;
 13. LexisNexis (<https://www.lexisnexis.com/>) helps to analyze court decisions and identify linguistic patterns of argumentation;
 14. Luminance (<https://www.luminance.com/>) is needed to process legal texts and highlight terms using machine learning;

15. Sonix (<https://sonix.ai>) automatically transcribes documents in audio and video formats (legal hearings, consultations, and interviews) into text, which facilitates effective information search and analysis.

Let us show how some of the AI tools listed can be used for research purposes. The authors checked the possibilities of the two AI applications – Sonix and ContractCrab because they provide different ways of dealing with legal texts.

Sonix is an automated service for converting audio and video files into text with support for more than 53 languages. In terms of the linguistic analysis of legal texts, this tool makes it possible to quickly and accurately transcribe the audio recordings of court sessions, interviews, conferences, and other speeches in audio and video formats, which significantly expands the possibilities of data collection and processing. Its built-in search, automatic summarization, and thematic and semantic analysis tools make it easier to identify key terms and phrases, which is especially important for analyzing legal texts in terms of their structure and content. Sonix also supports automatic translation, allowing users to work with multilingual legal documents.

The authors conducted the linguistic analysis of the text of the Court Hearing of the case Jacques v. Bondi 23-3015 (United States Court for the Ninth Circuit, 2025).

The audio recording is 38 minutes 57 seconds long.

The key features of Sonix and their description are presented below (Table 1).

Table 1. Features of the AI tool “Sonix” are useful for linguistic research.

Feature	Description	Example
Audio recording transcription	Provides basic text statistics	Words – 1603 Characters – 7978 Characters excluding spaces – 6376
	Is done within several seconds	The transcript needs polishing as words are transcribed with different levels of confidence: Very confident (89.27%) – 1298 words Fairly confident (10.04%) – 146 words Slightly confident (0.69%) – 10 words
	Provides speaker breakdown	Speaker 1 – 00:03:42 40.0% 618 words Speaker 2 – 00:03:34 38.6% 566 words Speaker 3 – 00:01:08 12.4% 245 words Speaker 4 – 00:00:50 9.0% 174 words
Transcript pagination	Allows one to split transcripts into small (10-minute) chunks; Provides better performance for longer files (Fig. 1)	
Timecodes	Realigns the individual words to the audio and regenerates each word’s time code. After a lot of editing, words and phrases may no longer align with the audio file	
Subtitles	Subtitles timeline can be shown or hidden	Each speaker is marked with a different color. The option “show timeline” splits the period of each speaker’s performance into “paragraphs” with marked time, one can play and pause each paragraph as many times as you need. The option “hide timeline” shows “paragraphs” marked with a different color according to the speaker
Vocabulary builder	Allows for the development of a custom dictionary	Words are added to a personal dictionary. The words are highlighted in the text and given a time mark: I plan to use about seven minutes for my argument and reserve the remainder for

Generate summary	Offers 2 options: transcript summary in a few sentences, transcript summary in a few bullet points	Result in a sentence format: Counsel for Mr. Jakes argued a due process violation because the BIA sent the briefing schedule to the wrong detention facility amid ICE transfer notice failures, and struggled to articulate specific prejudice beyond lack of notice while hinting at merits under Matter of NAM versus Matter of Y-L. The panel pressed on whether a motion to reopen was actually filed, the timeline of filings and service, prejudice requirements, and the aggravated felony framework, including whether a categorical analysis under Moncrieffe was ever conducted for the Utah drug conviction involving roughly 20–27 pounds of marijuana. Government counsel urged denial for lack of exhaustion or no prejudice, maintaining service was reasonably directed to Nye County at the time, asserting the conviction is an aggravated felony triggering Matter of Y-L’s particularly serious crime presumption, and arguing any presumption of prejudice would be rebutted by the BIA’s merits decision and the large-quantity drug-trafficking facts
Create chapters	Provides the functions: automatic theme detection, breakdown into logical sections, creating headings, and creating a table of contents	[00:00] Introduction and non-execution of a court order [01:20] The issue of prisoner transfer notifications [02:15] The question of the submitted petition and its status [04:30] Key issue: Violation of the right to due process [05:45] Discussion of the criterion of “Prejudice” [08:20] Discussion of legal standards: Matter of Y.L. vs. Matter of N-A-M [12:10] Defense argument: No drug possession [13:50] The question of the presumption of prejudice
Sentiment analysis	Allows for the uncovering of the emotions of the speakers	SEGMENT 1: PETITIONER’S COUNSEL; COURT 00:00:03 - 00:00:12 Sentiment: Neutral Rationale: Procedural opening with time allocation; respectful address to the court without emotional content. SEGMENT 2: COURT; PETITIONER’S COUNSEL 00:00:12 - 00:00:33 Sentiment: Negative Rationale: Court presses counsel on noncompliance with an order; tone is admonitory, indicating dissatisfaction. SEGMENT 3: PETITIONER’S COUNSEL 00:00:33 - 00:00:39 Sentiment: Negative Rationale: Brief, strained explanation “I have been in a medical issue” suggests defensiveness and tension.” SEGMENT 4: COURT; PETITIONER’S COUNSEL 00:00:39 - 00:00:47 Sentiment: Neutral Rationale: Court redirects to proceed; counsel attempts apology but is cut off; focus shifts to argument.
Thematic analysis	Allows for the revelation of dominant themes and patterns present in the transcript	1. Systemic incompetence and procedural failures: Manifestations in the text: failure to comply with court orders, loss of documents, violations by ICE; 2. Violation of the right to due Process Violation: Manifestations in the text: no notification, lack of opportunity to be heard; 3. The criterion of “Prejudice” as a procedural obstacle: Manifestations in the text: the requirement to prove damages, lack of substantive arguments, and weakness of the defense position; 4. Legal uncertainty and conflict of precedents: Manifestations in the text: the dilemma of Matter of Y.L. vs. Matter of N-A-M, different consequences; 5. Factual circumstances versus legal formality: Manifestations in the text: the defense argument, control issues of the court

Source: Compiled by the authors based on (Sonix, n.d.).

As shown in Figure 1, the transcript pagination feature divides the text into “pages.” The complete text is split into smaller parts, each of ten-minute length. Figure 1 shows the part of the text located on the first page (below the control panel on the left – “First Page”), and the

content of the next “page” of the text will be shown in the Next ten-minute page (below the control panel on the right – “Next Ten Minutes”). Another AI assistant, “ContractCrab,” provides several tools to analyze contracts. First, it is useful to identify the vocabulary

typical of legal English (Litvishko et al., 2025b) and its frequency, and the specifics of context use. Then, “ContractCrab” conducts a comparative analysis of legal documents as far as the style and linguistic changes are concerned, which is demonstrated below.

The authors carry out the research on the material of the text of “The Coca-Cola Company 2014 Equity Plan Stock Option Agreement” (Justia Business Contracts, 2014).

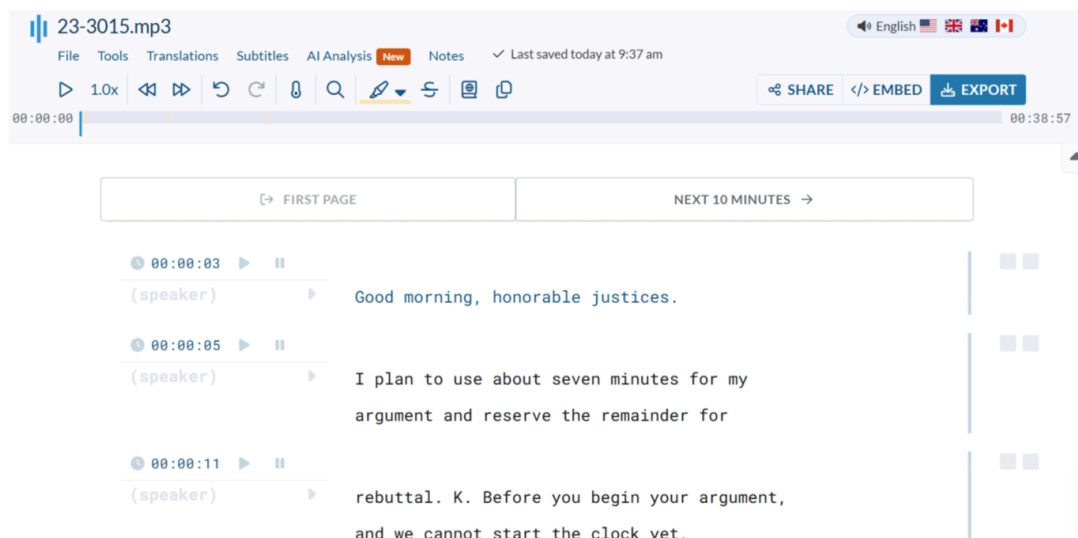


Figure 1. Transcript pagination performed by Sonix. *Source:* Compiled by the authors based on (Sonix, n.d.).

After the contract text is uploaded, the AI assistant highlights potentially ambiguous passages in purple. For each of these passages, ContractCrab suggests alternative wording to improve clarity. Given that the document dates back to 2014, legal language has evolved since

then. The revised wording proposed by the AI assistant reflects contemporary legal English, allowing us to observe how the language of law has changed over time (Table 2).

Table 2. Analysis of the contract by ContractCrab.

Old wording	New wording
Except as is otherwise explicitly provided in this Agreement and the Plan, non-vested Options are forfeited immediately following termination of employment for any reason, and vested Options expire the earlier of: a) six months following termination of employment for any reason, and b) the Expiration Date noted in the Option	Non-vested Options shall be forfeited immediately following termination of employment for any reason, except in cases of termination without cause, where a grace period of 90 days shall apply for the exercise of vested Options.

Source: Compiled by the authors based on (ContractCrab, n.d.).

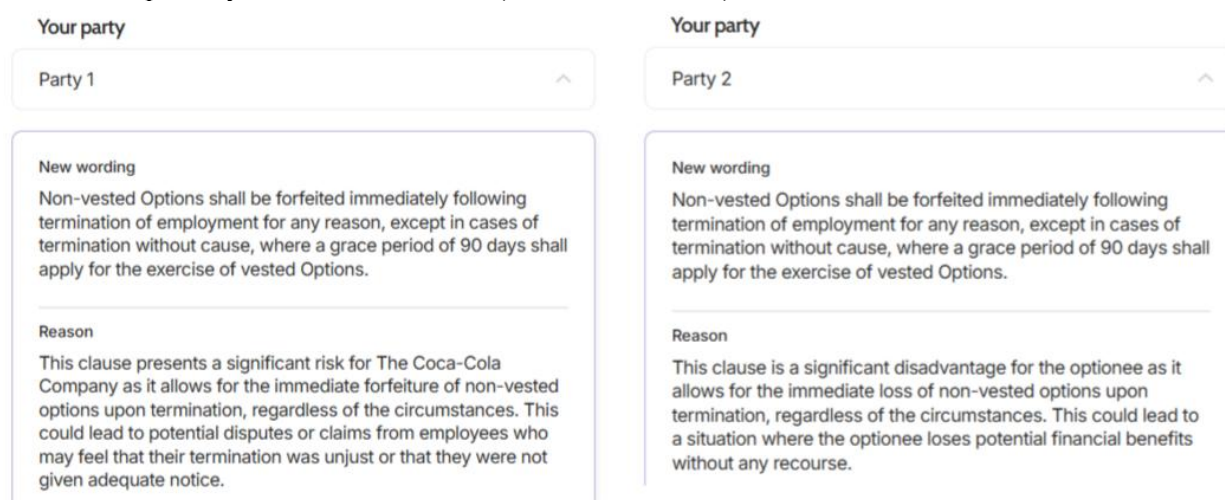


Figure 2. Comparison of the positions of both parties to the contract related to the new wording proposed by ContractCrab. *Source:* Compiled by the authors based on (ContractCrab, n.d.).

As can be clearly seen from the table, the language used to state legal norms has changed over time. This shift reflects the general trend toward simplification in English, which has developed under the influence of the Plain Language movement since the 1970s and aims to overcome the complexity of legal texts (Justia Business Contracts, 2014, p. 41). This conclusion is primarily relevant for linguists, as it highlights changes in the language itself.

Besides, the AI assistant provides a legally-grounded reason why new wording is better from the perspective of both parties (Figure 2).

Such an explanation of the reasons is useful for linguists, as it helps them understand how legal reasoning is expressed in language. It is also evident that the technologies described above facilitate the processing of large volumes of linguistically relevant information and significantly increase the accuracy of research and analysis.

4. DISCUSSION

The conducted research identified the key directions and features of the use of AI in the analysis of the legal English language. The analysis demonstrated that advanced AI tools significantly facilitate linguistic analysis by enabling the automation of terminology recognition, thematic and stylistic analysis, and the simplification of complex legal texts for a broad audience. Particular attention was given to the Sonix and ContractCrab platforms, which represent different approaches to processing audio and textual data and thereby expand the possibilities for collecting and analyzing linguistic information. Furthermore, an examination of ContractCrab's proposed reforms of legal

formulations revealed diachronic changes in legal language.

Nevertheless, despite significant progress, the integration of AI into legal linguistics remains limited for some reasons of both linguistic and extra-linguistic nature. On the one hand, legal texts are an extremely diverse corpus, including regulations, court decisions, contracts, commentaries, etc., each of which has its own specific structure, style, goals of communication, coupled with extremely difficult terminology. On the other hand, ethical issues of using AI pose risks to fundamental human rights and freedoms, including privacy, data protection, and non-discriminatory treatment (Dyadchenko & Kartashov, 2023; Hage, 2017). Automated analysis of legal texts can lead to errors and distortions, especially if the algorithms overlook the complexity and context of legal terminology, which can affect the legality and fairness of decisions.

5. CONCLUSION

Thus, the use of AI in studying legal English opens up new prospects for improving the effectiveness of linguistic analysis and practical work with legal texts. AI technologies not only enhance the automation of text processing but also expand opportunities for adapting materials to different levels of users' language proficiency. It is therefore important to implement AI tools widely while taking into account the specific features of legal discourse and its linguistic characteristics. A promising direction for further research is the deeper integration of AI into linguistic studies, aimed at developing integrated intelligent support systems and examining the impact of diachronic change on the evolution of legal language.

References:

- Armour, J., Parnham, R., & Sako, M. (2020). Unlocking the potential of AI for English law. *International Journal of the Legal Profession*, 28(1), 65-83. DOI: 10.1080/09695958.2020.1857765
- ContractCrab. (n.d.). *ContractCrab is an AI-powered platform designed to simplify the management of legal documents by providing concise summaries of lengthy contracts*. Retrieved from <https://www.trendingaitools.com/ai-tools/contractcrab/>; <https://contractcrab.com/> (Accessed 14 January 2026)
- Dyadchenko, A. A., & Kartashov, I. I. (2023). Ethical and legal aspects of artificial intelligence regulation. *Ius Publicum et Privatum: Online Scientific and Practical Journal of Private and Public Law*, 2(22), 34-44. DOI: 10.46741/2713-2811.2023.22.2.003
- Hage, J. (2017). Theoretical foundations for the responsibility of autonomous agents. *Artificial Intelligence and Law*, 25(3), 255-271. DOI: 10.1007/s10506-017-9208-7
- Justia Business Contracts. (2014). *The Coca-Cola Company 2014 equity plan stock option agreement with employee*. Retrieved from <https://contracts.justia.com/companies/coca-cola-309/contract/304655/> (Accessed 14 January 2026)
- Leonardi, V., & Giampieri, P. (2025). Artificial intelligence and the legal field. In *Legal English through the ages* (pp. 185-237). Cham, Switzerland: Palgrave Macmillan. DOI: 10.1007/978-3-032-02172-4_6
- Litvishko, O. M., Pronchenko, E. N., Konstantinova, A. S., & Miller, I. S. (2025a). AI tools in the analysis of legal discourse: Methods and prospects. In A. B. Karbekova, L. Žák, & I. Milenković (Eds.), *Big Data and Artificial Intelligence for decision-making in the smart economy* (pp. 283-292). Cham, Switzerland: Springer. DOI: 10.1007/978-3-031-78686-0_30

- Litvishko, O., Shiryaeva, T., Golubovskaya, E., & Mekeko, N. (2025b). Why English legal discourse is difficult to understand. *International Journal of Arabic-English Studies*, 25(1), 403-430. DOI: 10.33806/ijaes.v25i1.675
- Sonix. (n.d.). *Sonix is an AI-driven transcription service*. Retrieved from <https://www.trendingaitools.com/ai-tools/sonix/>; <https://sonix.ai> (Accessed 14 January 2026)
- Storey, V. C., Yue, W. T., Zhao, J. L., & Lukyanenko, R. (2025). Generative artificial intelligence: evolving technology, growing societal impact, and opportunities for information systems research. *Information Systems Frontiers*, 27, 2081-2102. DOI: 10.1007/s10796-025-10581-7
- Trending AI Tools. (n.d.). *Best AI Tools for Legal & Compliance*. Retrieved from <https://www.trendingaitools.com/ai-tools-category/legal-compliance/> (Accessed 14 January 2026)
- Tu, S. S., Cyphert, A., & Perl, S. J. (2024). Artificial intelligence: Legal reasoning, legal research and legal writing. *Minnesota Journal of Law, Science & Technology*, 25(2), 105-125. Retrieved from <https://scholarship.law.umn.edu/mjlst/vol25/iss2/11> (Accessed 14 January 2026)
- United States Court for the Ninth Circuit. (2025). *Case “Jacques v. Bondi 23-3015”*. Las Vegas, NV: US Court for the Ninth Circuit. Retrieved from <https://www.ca9.uscourts.gov/media/audio/?20251010/23-3015/> (Accessed 14 January 2026)

Olga B. Bagrintseva

Astrakhan State University named
after V.N. Tatishchev, Astrakhan,
Russia
bagrintsevaob@gmail.com
ORCID: 0000-0002-7117-0520

Olga M. Litvishko

Pyatigorsk State University,
Pyatigorsk, Russia
olitvishko@yandex.ru
ORCID: 0000-0001-8430-8501

Elena V. Ryumkova

Astrakhan State University named
after V.N. Tatishchev, Astrakhan,
Russia
khlopkova_elen@mail.ru
ORCID: 0009-0000-0484-1024
