UDC 81’342:004.85
DOI https://doi.org/10.26661/2414-1135-2026-103-20

Стаття поширюється на умовах ліцензії відкритого доступу CC BY 4.0

AI-ASSISTED ACOUSTIC ANALYSIS OF LEGAL DISCOURSE PROSODY: OPPORTUNITIES AND LIMITATIONS

Rudik I. V.
Candidate of Philological Sciences, Associate Professor,
Associate Professor at the Department of Foreign Languages
for Specific Purposes
Odesa I. I. Mechnikov National University
Vsevoloda Zmiienka str., 2, Odesa, Ukraine
orcid.org/0000-0001-6121-8060
i.rudik@onu.edu.ua

Ter-Grygorian M. H.
Candidate of Philological Sciences, Associate Professor,
Associate Professor at the Department of Foreign Languages
for Specific Purposes
Odesa I. I. Mechnikov National University
Vsevoloda Zmiienka str., 2, Odesa, Ukraine
orcid.org/0000-0001-7558-2533
maryna-ter-gryhorian@onu.edu.ua

Key words: *legal discourse, prosody, AI-assisted acoustic analysis, artificial intelligence, computational paralinguistics, forensic phonetics, speech processing, legal linguistics.*

The article examines the methodological potential of artificial intelligence (AI)-assisted acoustic analysis for investigating prosody in legal discourse. Prosody is viewed as an essential component of spoken legal communication, shaping the expression of epistemic stance, interpersonal relations, discourse organisation, and persuasive strategies through variations in pitch, intensity, rhythm, pausing, speech rate, and voice quality. The study addresses the limited attention paid to the acoustic dimension of legal discourse and evaluates the opportunities and challenges of integrating AI technologies into prosodic research. The object of the study is the prosodic organisation of spoken legal discourse; the subject is the methodological potential and limitations of AI-assisted acoustic analysis of legal speech. The aim of the study is to provide a criteria-based evaluation of AI-assisted acoustic analysis as a methodological framework for research on legal prosody. The research is theoretical and methodological, drawing on studies in acoustic phonetics, computational paralinguistics, forensic phonetics, legal linguistics, and speech processing. Descriptive, comparative, and critical-evaluative methods are applied to assess AI-assisted acoustic analysis according to six methodological criteria: validity, reliability, scalability, interpretability, ecological validity, and legal and ethical admissibility. The findings demonstrate that AI significantly improves the efficiency, reproducibility, and scalability of acoustic analysis by automating speech preprocessing, feature extraction, and computational modelling, enabling large-scale investigation of authentic legal speech. At the same time, important limitations remain, including reduced ecological validity, limited model transparency, algorithmic bias, risks of overinterpreting acoustic evidence, and legal and ethical concerns. The study argues that AI should complement rather than replace expert phonetic interpretation. It concludes that a human-in-the-loop approach integrating computational technologies with linguistic expertise provides the most reliable framework for future research on legal prosody while ensuring methodological transparency, scientific validity, and legal accountability.

ШТУЧНИЙ ІНТЕЛЕКТ У АКУСТИЧНОМУ АНАЛІЗІ ПРОСОДІЇ ЮРИДИЧНОГО ДИСКУРСУ: МОЖЛИВОСТІ ТА ОБМЕЖЕННЯ

Рудік І. В.
кандидат філологічних наук, доцент,
доцент кафедри іноземних мов професійного спрямування
Одеський національний університет імені І. І. Мечникова
вул. Всеволода Змієнка, 2, Одеса, Україна
orcid.org/0000-0001-6121-8060
i.rudik@onu.edu.ua

Тер-Григорьян М. Г.
кандидат філологічних наук, доцент,
доцент кафедри іноземних мов професійного спрямування
Одеський національний університет імені І. І. Мечникова
вул. Всеволода Змієнка, 2, Одеса, Україна
orcid.org/0000-0001-7558-2533
maryna-ter-gryhorian@onu.edu.ua

Ключові слова: *юридичний дискурс, просодія, штучний інтелект, AI-асистований акустичний аналіз, комп'ютерна паралінгвістика, судова фонетика, обробка мовлення, юридична лінгвістика.*

Статтю присвячено дослідженню методологічного потенціалу акустичного аналізу просодії юридичного дискурсу із застосуванням технологій штучного інтелекту. Просодію розглянуто як важливий компонент усного юридичного мовлення, що забезпечує реалізацію епістемічних, прагматичних, міжособистісних і дискурсивно-організаційних функцій через варіювання висоти основного тону, інтенсивності, темпу мовлення, ритму, паузації та якості голосу. Об'єктом дослідження обрано просодичну організацію усного юридичного мовлення, а його предметом є методологічний потенціал і обмеження AI-асистованого акустичного аналізу просодичних характеристик юридичного мовлення. Метою дослідження є оцінка можливостей і обмежень використання штучного інтелекту для акустичного аналізу просодичних характеристик юридичного мовлення. Теоретико-методологічною основою роботи стали дослідження з акустичної фонетики, комп'ютерної паралінгвістики, судової фонетики, юридичної лінгвістики та технологій обробки мовлення. У роботі використано описовий, порівняльний і критико-аналітичний методи для оцінювання AI-асистованого акустичного аналізу за критеріями валідності, надійності, масштабованості, інтерпретованості, екологічної валідності, а також юридичної й етичної прийнятності. Встановлено, що застосування технологій штучного інтелекту підвищує ефективність, відтворюваність і масштабованість акустичного аналізу завдяки автоматизації обробки мовлення, вилучення акустичних параметрів і комп'ютерного моделювання. Водночас виявлено обмеження, пов'язані з недостатньою прозорістю алгоритмів, можливими алгоритмічними упередженнями, ризиками надмірної інтерпретації акустичних даних, а також юридичними й етичними аспектами використання ШІ. Обґрунтовано доцільність застосування підходу human-in-the-loop, за якого технології штучного інтелекту доповнюють експертну фонетичну інтерпретацію. Зроблено висновок, що поєднання сучасних AI-технологій із лінгвістичною експертизою забезпечує наукову обґрунтованість і методологічну надійність досліджень просодії юридичного дискурсу.

Problem statement. Prosody constitutes a fundamental component of spoken communication, providing speakers with a wide range of acoustic resources for expressing interpersonal, pragmatic, and discourse-organising meanings. Through systematic variation in fundamental frequency (F0), intensity, temporal organisation, rhythm, pausing, and voice quality, speakers signal epistemic stance, emotional involvement, communicative intentions, and social relationships. Rather than serving as a mere complement to lexical and grammatical information, prosody actively shapes the interpretation of spoken interaction by influencing how utterances are understood within their communicative context [Korolova, Zhmayeva, 2021; Rudik, Vykhor, 2020]. The significance of prosody becomes particularly evident in legal discourse, where spoken interaction frequently determines the interpretation of evidence, the management of institutional roles, and the assessment of communicative credibility. Courtroom examination, judicial reasoning, advocacy, police interviews, mediation, and legal negotiation all rely not only on what participants say but also on how they say it. Variations in pitch, pausing, speech rate, or voice quality may contribute to the perception of confidence, hesitation, authority, cooperation, or resistance. Prosodic organisation is therefore an important dimension of legal communication, and its systematic investigation has implications for legal linguistics, forensic phonetics, and discourse analysis [Heffer, 2005; Akkurt, Korolova, Popova, 2021].

Despite its communicative importance, prosody remains one of the least explored dimensions of legal discourse. Previous research has predominantly concentrated on lexical, syntactic, rhetorical, and pragmatic characteristics of legal language, particularly within written legal texts [Heffer, 2005]. Although recent studies have increasingly examined spoken courtroom interaction, comprehensive acoustic investigations of prosody remain comparatively limited, largely because conventional phonetic analysis depends on time-consuming manual annotation and measurement. Software environments such as *Praat* provide highly accurate acoustic measurements but require extensive expert intervention throughout segmentation, annotation, and interpretation, thereby restricting both corpus size and the reproducibility of empirical findings [Boersma, Weenink, 2026].

Recent advances in artificial intelligence and speech technologies have substantially changed the methodology of acoustic phonetics. Automatic forced alignment, standardised acoustic feature extraction, computational paralinguistics, and deep-learning approaches now enable researchers to analyse large quantities of speech data with a level of efficiency and consistency that was previously unattainable. These developments have expanded the methodological possibilities for investigating prosodic variation across diverse communicative settings while facilitating the integration of quantitative speech analysis into broader linguistic research [McAuliffe et al., 2017; Eyben et al., 2016; Schuller, 2018].

Nevertheless, the increasing adoption of AI-based methods raises important methodological, legal, and ethical questions, particularly within the context of legal discourse. Most contemporary computational models have been developed and validated using acted emotional speech or general-purpose corpora that differ substantially from authentic legal communication. Courtroom interaction is spontaneous, strategically controlled, institutionally constrained, and frequently multilingual, making direct transfer of existing AI models methodologically problematic. Their application in legally sensitive contexts is further restricted by the limited interpretability of complex machine-learning systems, by demographic and accent-related biases, by uncertainty regarding ecological validity and by the absence of universally accepted standards for evidential reliability [Schuller, 2018; Koenecke et al., 2020; Morrison, 2009].

Although legal linguistics, forensic phonetics, and computational speech processing have each generated substantial bodies of research, these disciplines have evolved largely independently. Legal linguistic scholarship has provided detailed descriptions of institutional discourse but has devoted relatively little attention to acoustic realisation [Heffer, 2005]. Functional phonetic research has demonstrated the central role of prosody in expressing stance, interpersonal alignment, and information structure [Korolova, Zhmayeva, 2021; Rudik, Vykhor, 2020], whereas computational speech research has primarily focused on improving automatic recognition and classification algorithms [Eyben et al., 2016; Schuller, 2018]. A research gap therefore persists at the intersection of these disciplines.

The aim and objectives of the article. The aim of the study is to provide a criteria-based evaluation of the opportunities and limitations of AI-assisted acoustic analysis as a methodological framework for research on prosody in legal discourse.

The aim determines the following tasks: 1) to specify the communicative functions of prosody in legal discourse and to determine the acoustic parameters through which these functions are realised; 2) to systematise the AI technologies currently applied to the automatic processing and analysis of prosodic characteristics of speech; 3) to compare conventional manual acoustic analysis with AI-assisted analytical procedures in terms of measurement accuracy, reproducibility, scalability, and interpretability; 4) to evaluate AI-assisted acoustic analysis against six methodological criteria: validity, reliability, scalability, interpretability, ecological validity, and legal

and ethical admissibility; 5) to identify the methodological, technological, legal, and ethical constraints that restrict the application of AI to legally sensitive speech data; 6) to substantiate a human-in-the-loop model of research in which computational analysis is combined with expert phonetic interpretation.

The subject and object of the study. The object of the study is the prosodic organisation of spoken legal discourse, considered as a system of acoustic resources that participate in the construction of meaning in institutional legal communication.

The subject of the study is the methodological potential and the limitations of AI-assisted acoustic analysis as an instrument for investigating the prosodic parameters of legal speech.

Presentation of the main material. Materials and methods. The present research is theoretical and methodological in nature and does not involve the collection or experimental analysis of new speech data. Instead, it synthesises and critically evaluates contemporary research in acoustic phonetics, computational paralinguistics, forensic phonetics, legal linguistics, and artificial intelligence in speech processing [Heffer, 2005; Schuller, 2018]. The analytical framework is based on two complementary sources of evidence. The first encompasses methodological studies devoted to automatic speech processing, acoustic feature extraction, machine learning, and computational modelling of prosodic phenomena [Eyben et al., 2016; McAuliffe et al., 2017]. The second consists of linguistic research describing the communicative functions of prosody in legal discourse, including courtroom examination, judicial speech, advocacy, police interviews, mediation, and legal negotiation [Akkurt, Korolova, Popova, 2021; Heffer, 2005].

Methodologically, the study combines descriptive, comparative, and critical-evaluative approaches.

The descriptive-analytical method is employed to identify and systematise the principal prosodic parameters that contribute to meaning construction in legal discourse, as well as the AI-based technologies currently available for their automatic analysis [Korolova, Zhmayeva, 2021; Rudik, Vykhor, 2020].

The comparative method is used to examine the respective advantages and limitations of conventional manual acoustic analysis and AI-assisted analytical procedures across shared methodological dimensions, including measurement accuracy, reproducibility, scalability, and interpretability [Boersma, Weenink, 2026; Schuller, 2018].

The critical-evaluative method constitutes the principal analytical framework of the study. AI-assisted acoustic analysis is assessed against six interrelated methodological criteria: validity, reliability, scalability, interpretability, ecological validity, and legal and ethical admissibility [Morrison, 2009; Koenecke et al., 2020].

The analysis considers the principal acoustic dimensions of prosody, including melodic parameters (fundamental frequency, pitch range, contour configuration, and pitch movement), dynamic characteristics (intensity and loudness variation), temporal organisation (segmental duration, speech rate, articulation rate, and pausing), rhythmic structure, and phonatory features such as creaky voice, breathiness, and vocal tension [Eyben et al., 2016].

Particular attention is devoted to the relationship between acoustic measurability and functional interpretation. While parameters such as fundamental frequency, duration, and intensity can be extracted automatically with a high degree of precision, their communicative significance cannot be inferred directly from numerical measurements. Conversely, the phenomena that carry the greatest interactional weight, among them subtle variations in voice quality and rhythm, remain difficult to model computationally [Schuller, 2018].

To conceptualise contemporary AI-assisted acoustic analysis, the study adopts a four-stage analytical pipeline. The first stage comprises speech preprocessing and automatic forced alignment, whereby recordings are segmented and temporally synchronised with orthographic transcripts [McAuliffe et al., 2017]. The second stage involves standardised acoustic feature extraction using internationally recognised parameter sets [Eyben et al., 2016]. During the third stage, supervised and self-supervised machine-learning models establish relationships between acoustic representations and communicative categories such as emotion, stance, or speaker identity [Schuller, 2018]. Finally, the fourth stage consists of linguistic interpretation, during which computational outputs are critically evaluated in relation to discourse context, communicative function, and legal relevance [Morrison, 2009].

Results and discussion. *AI-assisted acoustic analysis as a methodological paradigm for legal prosody research.* The development of artificial intelligence has reshaped methodological approaches to speech analysis, shifting acoustic phonetics from predominantly manual investigation towards large-scale computational analysis. In legal discourse research, this transformation is particularly significant because prosodic phenomena frequently extend beyond isolated linguistic observations and require the examination of extensive speech corpora representing diverse communicative situations. AI-assisted acoustic analysis therefore operates as a methodological paradigm that expands both the scale and the analytical scope of legal prosody research [Schuller, 2018; Eyben et al., 2016].

Unlike conventional phonetic analysis, which relies heavily on manual segmentation and annotation, contemporary AI-assisted workflows automate

the majority of repetitive analytical procedures while maintaining high levels of measurement consistency. Automatic forced alignment, standardised extraction of acoustic parameters, and machine-learning-based classification considerably reduce the time required to process speech recordings, enabling the analysis of complete courtroom proceedings, mediation sessions, police interviews, or negotiation corpora rather than isolated illustrative excerpts. This transition from small datasets to large, systematically annotated corpora represents a qualitative methodological shift, facilitating statistically robust comparisons across speakers, institutional roles, communicative genres, and linguistic communities [McAuliffe et al., 2017; Eyben et al., 2016].

An equally important advantage concerns methodological reproducibility. Manual acoustic analysis inevitably incorporates a degree of researcher-dependent variation, particularly during segmentation, annotation, and parameter selection. By contrast, AI-assisted workflows employ standardised computational procedures that can be replicated across laboratories and research projects. The use of internationally recognised acoustic parameter sets, such as GeMAPS, promotes methodological transparency and facilitates direct comparison between independent studies. Such reproducibility is particularly valuable within forensic and legal-linguistic research, where analytical procedures must withstand expert scrutiny and, potentially, judicial evaluation [Eyben et al., 2016; Morrison, 2009].

Beyond improvements in efficiency and reproducibility, artificial intelligence considerably enhances the analytical sensitivity of acoustic research. Prosodic meaning rarely emerges from isolated acoustic parameters; rather, it reflects complex interactions among pitch movement, temporal organisation, intensity, rhythm, and voice quality. Machine-learning algorithms, particularly deep neural and self-supervised architectures, are capable of modelling these multidimensional relationships simultaneously, thereby identifying subtle acoustic regularities that often remain inaccessible through auditory perception or conventional statistical analysis. AI-assisted approaches therefore open new opportunities for investigating communicative phenomena such as epistemic stance, interpersonal dominance, credibility, uncertainty, and strategic alignment in institutional interaction [Schuller, 2018].

The methodological value of AI extends further through its capacity to support comparative and longitudinal research. Standardised computational pipelines allow identical analytical procedures to be applied across different legal systems, languages, institutional settings, and communicative contexts. Such consistency facilitates cross-cultural comparisons of prosodic behaviour while enabling researchers to examine changes in communicative strategies across different stages of legal proceedings. Longitudinal analyses of individual speakers, institutional practices, or negotiation processes become methodologically feasible only when speech processing is sufficiently automated to accommodate large volumes of data [McAuliffe et al., 2017; Schuller, 2018].

Rather than replacing established phonetic methodology, AI significantly broadens the empirical possibilities available to legal discourse research. Computational analysis should therefore be understood as an extension of traditional instrumental phonetics, combining large-scale quantitative processing with linguistically informed interpretation. Within this perspective, artificial intelligence functions primarily as an analytical instrument that increases methodological precision, consistency, and scalability while leaving theoretical interpretation and contextual evaluation to the researcher [Boersma, Weenink, 2026; Morrison, 2009].

Functional applications of AI-assisted prosodic analysis in legal communication. The practical value of AI-assisted acoustic analysis becomes evident when prosodic measurements are interpreted within the communicative functions that characterise legal discourse. Prosody is an essential resource for organising institutional interaction, negotiating interpersonal relationships and constructing legal argumentation. Computational analysis therefore acquires scientific value only when acoustic parameters are systematically related to communicative function [Korolova, Zhmayeva, 2021; Akkurt, Korolova, Popova, 2021].

From a functional perspective, four principal dimensions of legal communication can be distinguished. The *epistemic dimension* concerns the expression of certainty, uncertainty, commitment, and evidential stance. Variations in pitch movement, pausing behaviour and speech rate frequently accompany communicative strategies through which speakers signal confidence, reservation, or incomplete knowledge. AI-assisted analysis enables these prosodic configurations to be examined quantitatively across extensive datasets, thereby providing empirical evidence for linguistic theories of stance construction [Korolova, Zhmayeva, 2021].

The *relational dimension* reflects the negotiation of institutional authority and interpersonal alignment. Courtroom interaction is inherently asymmetrical, involving judges, lawyers, witnesses, defendants, and other participants whose institutional roles shape conversational dynamics. Acoustic indicators such as pitch range, loudness variation, interruption timing, and speech rhythm contribute to the construction of dominance, deference, cooperation, and procedural control. Automated analysis enables these interactional patterns to be investigated objectively across

multiple communicative contexts [Rudik, Vykhor, 2020; Heffer, 2005].

The *affective dimension* encompasses the controlled expression and regulation of emotion. Unlike everyday conversation, legal communication generally requires speakers to regulate emotional display according to institutional expectations. Computational analysis of prosodic variation provides opportunities to investigate how emotional expression is strategically moderated during courtroom testimony, police interviews, mediation, and legal negotiation. Nevertheless, such analyses should remain descriptive rather than diagnostic, since current scientific evidence does not support direct inference of psychological states from acoustic correlates alone [Schuller, 2018; Morrison, 2009].

The *rhetorical dimension* concerns persuasive communication. Judicial reasoning, advocacy, and legal argumentation rely extensively on prosodic foregrounding, strategic emphasis, temporal organisation, and intonational structuring to enhance communicative effectiveness. AI-assisted acoustic analysis permits these persuasive strategies to be examined quantitatively across extended discourse, thereby complementing traditional rhetorical and discourse-analytic approaches [Akkurt, Korolova, Popova, 2021; Heffer, 2005].

Collectively, these four dimensions establish a theoretically grounded framework through which computational acoustic measurements can be interpreted within legal communication. Rather than treating prosodic parameters as isolated numerical variables, this framework connects acoustic evidence with broader communicative functions, thereby strengthening the explanatory value of AI-assisted analysis [Korolova, Zhmayeva, 2021; Akkurt, Korolova, Popova, 2021].

Methodological, legal and ethical limitations of AI-assisted prosodic analysis. Despite the considerable methodological advantages offered by artificial intelligence, its application to the acoustic analysis of legal discourse remains constrained by a number of interrelated methodological, technological, legal, and ethical limitations. These challenges should not be regarded as isolated shortcomings of individual algorithms but rather as structural characteristics of the current stage of AI development. The scientific value of AI-assisted prosodic analysis therefore depends as much on how explicitly these limitations are recognised and methodologically controlled as on technological performance [Schuller, 2018; Morrison, 2009].

The most fundamental limitation concerns ecological validity. Contemporary machine-learning models are typically trained on speech corpora collected under highly controlled experimental conditions, frequently involving acted emotional expressions, studio-quality recordings, or conversational speech that differs substantially from authentic legal interaction. By contrast, courtroom proceedings, police interviews, mediation sessions, and legal negotiations are characterised by spontaneous speech, institutional constraints, strategic self-monitoring, overlapping dialogue, varying recording quality, and considerable acoustic variability. As a result, computational models demonstrating high performance on benchmark datasets often exhibit substantially lower reliability when applied to naturally occurring legal discourse. The transferability of AI systems across communicative domains, therefore, remains one of the principal methodological challenges in legal speech research [Schuller, 2018; McAuliffe et al., 2017].

A closely related limitation concerns *model interpretability*. Recent advances in deep learning have produced highly accurate computational models capable of identifying complex multidimensional acoustic patterns. However, improvements in predictive performance have frequently been accompanied by a reduction in model transparency. Deep neural networks generally provide limited insight into the acoustic evidence underlying their predictions, thereby complicating scientific interpretation and reducing analytical accountability. While such opacity may be acceptable in engineering applications focused primarily on prediction accuracy, it represents a significant obstacle in legal contexts, where expert conclusions must be transparent, reproducible, and open to independent examination. Explainability should therefore be treated as a methodological requirement rather than an optional feature of AI-assisted legal analysis [Schuller, 2018; Morrison, 2009].

Another significant challenge concerns algorithmic bias and demographic fairness. Koenecke and colleagues showed that speech technologies perform unequally across demographic groups, including speakers who differ in age, gender, regional accent, first language or ethnic background [Koenecke et al., 2020]. These disparities largely reflect imbalances within training datasets rather than intrinsic linguistic differences. Nevertheless, when AI systems are applied to legal communication, such biases may produce serious practical consequences. Prosodic characteristics associated with second-language speech, regional pronunciation, or culturally specific interactional styles may be erroneously interpreted as indicators of uncertainty, emotional instability, or communicative unreliability. Such systematic distortions directly conflict with the legal principles of equality, procedural fairness, and non-discrimination.

Equally important is the problem of *overinterpretation.* Although computational models can successfully identify statistical relationships between acoustic parameters and communicative categories, such associations should not be interpreted as direct evidence of underlying psychological or cog-

nitive states. Prosodic manifestations of hesitation, increased speech rate, vocal tension, or pitch variability are inherently multifunctional and context-dependent. They may reflect cognitive load, communicative strategy, emotional arousal, cultural convention, interactional accommodation, or numerous other factors unrelated to deception or credibility. Current scientific evidence does not support reliable inference of truthfulness, deception, or internal mental states solely on the basis of acoustic information. AI-generated predictions should therefore be read as probabilistic analytical indicators rather than diagnostic conclusions [Schuller, 2018; Morrison, 2009].

Beyond methodological considerations, the application of AI to legal speech analysis raises a series of *legal and ethical concerns*. The collection, storage, and computational processing of spoken legal data inevitably involve issues of privacy protection, informed consent, confidentiality, and data governance. Additional challenges arise when AI-generated acoustic evidence is introduced within forensic or judicial contexts, where standards of admissibility require demonstrated scientific validity, methodological reliability, and transparent documentation of analytical procedures. Compliance with these requirements cannot be guaranteed through technological performance alone but depends equally on appropriate institutional regulation and expert oversight [Morrison, 2009; Koenecke et al., 2020].

Taken together, these limitations demonstrate that the principal challenge facing AI-assisted legal phonetics is not the replacement of human expertise but its responsible integration with computational technologies. The central methodological question is therefore not whether artificial intelligence should be employed in legal discourse research, but under which conditions its application can be considered scientifically valid, ethically responsible, and legally defensible [Morrison, 2009; Schuller, 2018].

Conclusions and prospects for further research. The preceding analysis suggests that neither complete reliance on automated systems nor exclusive dependence on traditional manual analysis represents an adequate methodological solution. Instead, the strengths of both approaches can be integrated within a complementary analytical framework in which artificial intelligence enhances efficiency and consistency, while expert phonetic interpretation ensures theoretical validity, contextual sensitivity, and professional accountability. Within such a framework, AI performs those analytical tasks for which computational methods are demonstrably superior, including automatic segmentation, acoustic feature extraction, large-scale corpus processing, statistical modelling, and the identification of multidimensional acoustic regularities. Human experts, by contrast, remain responsible for linguistic interpretation, evaluation of communicative context, verification of computational outputs, and formulation of scientifically justified conclusions. Rather than functioning as an autonomous decision-maker, artificial intelligence serves as an advanced analytical instrument supporting expert judgement.

This human-in-the-loop paradigm also provides practical mechanisms for addressing the limitations identified above. Ecological validity can be strengthened through the construction of representative legal speech corpora containing authentic recordings collected under realistic institutional conditions. Interpretability may be improved through the adoption of explainable AI techniques and transparent reporting of modelling procedures. Algorithmic bias requires systematic validation across demographic subgroups and continuous monitoring of model performance within diverse linguistic communities. Finally, legal and ethical compliance demands standardised procedures governing data protection, informed consent, documentation of analytical workflows, and independent expert verification.

The future development of AI-assisted legal phonetics should therefore be guided not only by advances in computational performance but also by rigorous methodological standards ensuring that technological innovation remains compatible with the principles of scientific transparency, legal accountability, and ethical responsibility. Under these conditions artificial intelligence can become a standard component of legal discourse research without displacing human expertise in the interpretation of linguistic evidence.

BIBLIOGRAPHY

1. Рудік І. В., Вихор О. Ф. Просодичні маркери соціального статусу мовця. Записки з романо-германської філології. 2020. Вип. 1 (44). С. 273–279. DOI: https://doi.org/10.18524/2307-4604.2020.1(44).211020.
2. Akkurt V., Korolova T., Popova O. Persuasion Prosody in Prosecutor's Speech: Ukrainian and English. Revista Amazonia Investiga. 2021. Vol. 10, no. 47. P. 125–141. DOI: https://doi.org/10.34069/ai/2021.47.11.13.
3. Boersma P., Weenink D. Praat: Doing Phonetics by Computer [Computer program]. Version 7.0.02. 2026. URL: https://www.praat.org (дата звернення: 05.09.2026).
4. Eyben F., Scherer K. R., Schuller B. W. et al. The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. IEEE Transactions on Affective Computing. 2016. Vol. 7, no. 2. P. 190–202. DOI: https://doi.org/10.1109/TAFFC.2015.2457417.
5. Heffer C. The Language of Jury Trial: A Corpus-Aided Analysis of Legal-Lay Discourse.

London : Palgrave Macmillan, 2005. 253 p. DOI: https://doi.org/10.1057/9780230502888.
6. Koenecke A., Nam A., Lake E. et al. Racial Disparities in Automated Speech Recognition. Proceedings of the National Academy of Sciences. 2020. Vol. 117, no. 14. P. 7684–7689. DOI: https://doi.org/10.1073/pnas.1915768117.
7. Korolova T., Zhmayeva N. Prosody of Negative Modality in Translation. Науковий вісник Південноукраїнського національного педагогічного університету імені К. Д. Ушинського: Лінгвістичні науки. 2021. № 32. С. 74–83. DOI: https://doi.org/10.24195/2616-5317-2021-32-6.
8. McAuliffe M., Socolof M., Mihuc S. et al. Montreal Forced Aligner: Trainable Text-Speech Alignment Using Kaldi. Proceedings of Interspeech 2017. Stockholm, 2017. P. 498–502. DOI: https://doi.org/10.21437/Interspeech.2017-1386.
9. Morrison G. S. Forensic Voice Comparison and the Paradigm Shift. Science & Justice. 2009. Vol. 49, no. 4. P. 298–308. DOI: https://doi.org/10.1016/j.scijus.2009.09.002.
10. Schuller B. W. Speech Emotion Recognition: Two Decades in a Nutshell, Benchmarks, and Ongoing Trends. Communications of the ACM. 2018. Vol. 61, no. 5. P. 90–99. DOI: https://doi.org/10.1145/3129340.

REFERENCES

1. Rudik, I. V., & Vykhor, O. F. (2020). Prosodychni markery sotsialnoho statusu movtsia [Prosodic markers of the speaker's social status]. Zapysky z Romano-Hermanskoi Filolohii, 1 (44), 273–279. https://doi.org/10.18524/2307-4604.2020.1(44).211020
2. Akkurt, V., Korolova, T., & Popova, O. (2021). Persuasion prosody in prosecutor's speech: Ukrainian and English. Revista Amazonia Investiga, 10(47), 125–141. https://doi.org/10.34069/ai/2021.47.11.13
3. Boersma, P., & Weenink, D. (2026). Praat: Doing phonetics by computer (Version 7.0.02) [Computer software]. https://www.praat.org
4. Eyben, F., Scherer, K. R., Schuller, B. W., Sundberg, J., André, E., Busso, C., Devillers, L. Y., Epps, J., Laukka, P., Narayanan, S. S., & Truong, K. P. (2016). The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for voice research and affective computing. IEEE Transactions on Affective Computing, 7(2), 190–202. https://doi.org/10.1109/TAFFC.2015.2457417
5. Heffer, C. (2005). The language of jury trial: A corpus-aided analysis of legal-lay discourse. Palgrave Macmillan. https://doi.org/10.1057/9780230502888
6. Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. (2020). Racial disparities in automated speech recognition. Proceedings of the National Academy of Sciences, 117(14), 7684–7689. https://doi.org/10.1073/pnas.1915768117
7. Korolova, T., & Zhmayeva, N. (2021). Prosody of negative modality in translation. Naukovyi Visnyk of South Ukrainian National Pedagogical University Named after K. D. Ushynsky: Linguistic Sciences, (32), 74–83. https://doi.org/10.24195/2616-5317-2021-32-6
8. McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: Trainable text-speech alignment using Kaldi. In Proceedings of Interspeech 2017 (pp. 498–502). ISCA. https://doi.org/10.21437/Interspeech.2017-1386
9. Morrison, G. S. (2009). Forensic voice comparison and the paradigm shift. Science & Justice, 49(4), 298–308. https://doi.org/10.1016/j.scijus.2009.09.002
10. Schuller, B. W. (2018). Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends. Communications of the ACM, 61(5), 90–99. https://doi.org/10.1145/3129340

Дата першого надходження статті до видання: 20.07.2026
Дата прийняття статті до друку після рецензування: 23.08.2026
Дата публікації (оприлюднення) статті: 18.11.2026