



Научно-исследовательский журнал «Вестник филологических наук / Philological Sciences Bulletin»
<https://vfn-journal.ru>
2025, Том 5, № 6 / 2025, Vol. 5, Iss. 6 <https://vfn-journal.ru/archives/category/publications>
Научная статья / Original article
Шифр научной специальности: 5.9.8. Теоретическая, прикладная и сравнительно-сопоставительная лингвистика (филологические науки)
УДК 81.139

¹ Макович Г.В.

¹ Уральский юридический институт МВД России

Идиолектные маркеры в дипфейках как метод идентификации субъекта контента в юридической практике

Аннотация: цель статьи – разработать метод идентификации автора дипфейк-контента на основе идиолектных маркеров для применения в юридической практике. Актуальность проблемы для юридической практики и теории очень высокая: дипфейки создают новые риски для достоверности доказательств, нарушают права личности, поэтому нужны методы их надежной идентификации, такие как анализ идиолектных маркеров. Точность и операционность методики важны для верных правовых решений в условиях цифровизации.

В статье представлены методика разработки многоуровневенного идиолектного профиля субъекта речи и алгоритм действий по установлению вероятного авторства текста, необходимых для верных правовых решений в условиях цифровизации. В методике показаны маркеры идиолекта на фонетико-орфографическом, морфологическом, лексико-семантическом, синтаксическом, pragматическом уровнях. Излагается алгоритм действий в юридической практике при установлении идиолектного профиля. Методика многоуровневенного анализа идиолектного профиля субъекта речи может использоваться в судебной лингвистике для атрибуции любых анонимных текстов, установления plagiat или выявления фальсификаций. Проблемой в реализации многоуровневенного анализа идиолектного профиля является индивидуальная вариативность языковых единиц в идиолекте, индивидуальный вариант языка. В статье обосновано, что разработка методики анализа идиолектного профиля требует решения проблемы определения границ вариативности языковых признаков идиолекта.

Ключевые слова: дипфейк, идиолектные маркеры, многоуровневый идиолектный профиль, идиолектная систематизация, авторская атрибуция, юрислингвистика

Для цитирования: Макович Г.В. Идиолектные маркеры в дипфейках как метод идентификации субъекта контента в юридической практике // Вестник филологических наук. 2025. Том 5. № 6. С. 71 – 75.

Поступила в редакцию: 13 мая 2025 г.; Одобрена после рецензирования: 01 июня 2025 г.; Принята к публикации: 16 июня 2025 г.

¹ Makovich G.V.

¹ Ural Law Institute of the Ministry of Internal Affairs of Russia

Idiolect markers in deepfakes as a method of identifying the subject of content in legal practice

Abstract: the purpose of the article is to develop a method for identifying the author of deepfake content based on idiolect markers for use in legal practice. The relevance of the problem for legal practice and theory is very high: deepfakes create new risks for the reliability of evidence, violate individual rights, therefore methods of their reliable identification are needed, such as the analysis of idiolect markers. The accuracy and operability of the methodology are important for correct legal decisions in the context of digitalization.

The article presents a methodology for developing a multi-level idiolect profile of the subject of speech and an algorithm of actions to establish the probable authorship of the text, necessary for correct legal decisions

in the context of digitalization. The methodology shows idiolect markers at phonetic-orthographic, morphological, lexical-semantic, syntactic, and pragmatic levels. The algorithm of actions in legal practice when establishing an idiolect profile is described. The method of multilevel analysis of the idiolect profile of the subject of speech can be used in forensic linguistics to attribute any anonymous texts, establish plagiarism or identify falsifications, the problem in implementing a multilevel analysis of the idiolect profile is the individual variability of linguistic units in the idiolect, an individual language variant. The article substantiates that the development of a methodology for analyzing the idiolect profile requires solving the problem of determining the boundaries of variation in idiolectical features.

Keywords: deepfake, idiolect markers, multilevel idiolect profile, idiolect systematization, author's attribution, legal linguistics

For citation: Makovich G.V. Idiolect markers in deepfakes as a method of identifying the subject of content in legal practice. Philological Sciences Bulletin. 2025. 5 (6). P. 71 – 75.

The article was submitted: May 13, 2025; Approved after reviewing: June 01, 2025; Accepted for publication: June 16, 2025.

Введение

Технологии искусственного интеллекта стремительно изменяют современную коммуникацию и создают большое количество правовых коллизий, проявляя недостаточную проработанность правовой базы, не позволяя установить эффективный правовой контроль над новыми видами коммуникативной деятельности в цифровой среде, над деструктивным контентом, создающим угрозы общественному развитию.

Новые риски для юридической практики создают сгенерированные тексты различных видов. К их числу относятся дипфейки. Классический дипфейк представляет собой синтетический медиатекст (видео, аудио, фото), созданный с помощью искусственного интеллекта, обычно нейросетей, для имитации реальных людей или событий, с целью развлечения, дискредитации, дезинформации, мошенничества, обмана и прочих неблаговидных целей неправового характера [4]. Дипфейки нарушают права личности, такие как право на имя – неотчуждаемое субъективное право, являющееся личным неимущественным правом человека, право на частную жизнь, честь, достоинство и деловую репутацию. Дипфейки могут распространять дезинформацию о том или ином субъекте путём создания фейковых видео или аудио, где человек, обычно публичная личность, якобы произносит оскорбительные высказывания или делает то, чего не совершил, например, компрометирующие поступки. Дипфейки могут быть связаны с распространением ложных новостей, destabilизирующих общество. Дипфейки могут быть связаны с мошенничеством, а именно: с вымогательством с помощью видео или аудиосообщений от имени знакомых, родственников, руководителей.

Дипфейки могут вводить правоохранительные органы и суд в заблуждение, поэтому нужны методы их надежной идентификации. Таким методом может стать анализ идиолектных маркеров. Разработка метода анализа идиолектных маркеров важна для теории права, поскольку представляет собой развития правовых подходов к оценке цифровых доказательств для защиты прав участников судебного процесса в условиях цифровизации.

Цель статьи – разработать метод идентификации автора дипфейк-контента на основе идиолектных маркеров для применения в юридической практике.

Объект исследования – текст как компонент дипфейк-контента, предмет – идиолектные маркеры, под которыми в статье понимаются индивидуальные языковые особенности, используемые для авторской атрибуции синтетических материалов.

Материалы и методы исследований

В работе применяются общетеоретический метод исследования, метод лингвистического экспериментного анализа. Ключевой особенностью методики анализа идиолекта является сопоставление нескольких текстов одного субъекта речи и разных субъектов речи. Используются возможности количественных методов для моделирования индивидуальных особенностей языкового стиля говорящего, что связано с всё возрастающими потребностями объективации в анализе речи и с поиском формализованных способов выявления индивидуальных признаков речи. В качестве количественных методик используют автоматизированную стилеметрию, а также ручной или полуручной лингвистический анализ с опорой на устойчивые идиолектные маркеры.

Термин идиолект относится к области исследования индивидуально-языковых характеристик человека. Проблемами выявления индивидуальных особенностей языка занимается антропологическая лингвистика, главным объектом которой является человеческий фактор в языке, особенности речи отдельной личности.

В основном антропологическая лингвистика изучает речь творческих и публичных медиаличностей, профессионально владеющих языком, – как письменную, так и устную. Однако развитие интернет-коммуникации привело к появлению большого объема текстов рядовых носителей языка, чьё авторство требуется идентифицировать [8]. Это расширяет предметное поле в аспекте практического анализа идиолекта и теоретической разработки проблемы. У исследователей расширился объем и видовое разнообразие языкового материала, требующего теоретического осмысления. Также возникла насущная общественная потребность на надежные методики идентификации автора анонимного текста, содержащего противоправный контент или сконструированного с противоправной целью.

Термин идиолект употребляется в одном ряду с терминами идиостиль, языковая личность, речевой портрет языковой личности, но каждый из них имеет предметную специфику, хотя в целом все они связаны с исследованием индивидуального языка. Под идиолектом принято понимать весь комплекс формальных и стилистических характеристик языка индивида, который создаётся отбором и предпочтениями языковых единиц, использованием их сочетаемости и композиционных возможностей, творческими речевыми действиями [7]. Идиолект складывается на основе личного когнитивного и коммуникативного опыта субъекта речи, в его взаимодействии с социальными и речевыми эталонами, нормами.

Идиолект является лингвистической константой, составляющей основу высказываний индивида [3]. Иногда понятие идиолект трактуют широко и включают в него весь комплекс фонологических, лексических, синтаксических и стилистических характеристик речи индивида, рассматривая в составе идиолекта и фонетических особенностей устной речи конкретного носителя (тепп речи, интонация, ударения, голосовой диапазон), и особенности орфографии и пунктуации в его письменной речи, и особенности речи на лексико-синтаксическом и стилистическом уровнях [2]. При изучении идиолекта используют методики актуализированного лексикона, актуализированного синтаксиса, представляя вербальную и структурную сеть индивидуальной речи говорящего, доминанты его языкового сознания.

Идиостелевым приёмом (признаком идиолекта) не является использование частотной общеупотребительной лексики, типовых синтаксических конструкций, и с помощью статистических методов необходимо отграничить индивидуальные речевые явления в текстах субъекта речи и общезначимые явления. В связи с этим необходимо выявлять индивидуальную авторскую частотную лексику и текстовые лексические универсалии [6]. Идиолектная систематизация лексики текстов одного субъекта позволяет выстроить понятийную, смысловую иерархию, которая определяется структурой его языковой картины мира.

Результаты и обсуждения

Наиболее разработанными на сегодня являются методики работы с лексико-семантическим уровнем текста субъекта речи. Выявляются индивидуально-частотные слова, выражения, употребительность выразительных средств языка. Теоретически значимыми являются разработка и уточнение базовых терминов: концептуальная система автора, индивидуально-авторская частотная лексика, текстовые лексические универсалии, а также синтагматический профиль автора с учётом синтагматической активности слов.

Однако необходим также анализ текста на уровнях фонетико-орфографическом, морфологическом, синтаксическом, прагматическом, в результате чего выявляются особенности произношения и написания, грамматические предпочтения, предпочтения в выборе коммуникативных целей, стратегий и тактик. Только в ходе комплексного анализа возможно выявить многоуровневый идиолектный профиль субъекта речи.

На фонетико-орфографическом уровне можно выделить ряд маркеров идиолекта. Один из маркеров отражает региональные черты личности, такие как место рождения и место постоянного проживания. Региональные маркеры субъекта речи являются доминирующими, и в этом смысле они первичны по отношению к другому маркеру, проявляющему социальные характеристики субъекта речи, такие как принадлежность к определенному социальному слою, возрасту, профессии и др. Принадлежность к определенному социальному классу как ограниченной группе людей, наряду с включенностью в определенную социальную сеть с иерархией связей внутри и вне ее, оказывают значительное влияние на формирование типа речевого поведения, в том числе, на тип произношения [1].

Следующим диагностическим маркером идиостиля являются интонационные характеристики личности: темп речи, ее мелодика, изменения тона в пределах интонационной конструкции, в своеобразии сочетаний интонационных конструкций, в способе паузирования и выделения слов, которые несут смысловую и экспрессивную нагрузку. Существуют речевые пристрастия в выборе орфоэпического варианта, в фонетическом эллипсисе, в простоте или сложности интонационной модели, в особенностях употребления экспрессивной интонации. Маркером идиолекта на фонетико-орфографическом уровне являются опечатки и устойчивые ошибки (пропуск мягкого знака, смешение "о / а" в безударных позициях и многое другое).

К маркерам идиолекта на морфологическом уровне относятся предпочтения определенных грамматических форм (например, видовых, временных форм, причастий, деепричастий), их ненормативное образование. Распространены методика анализа частотности употребления частей речи и грамматических форм и методика последовательности употребления частей речи [5]. Эти методики показали продуктивность при решении вопроса о идиолектной принадлежности текста.

Синтаксические особенности речи выступают в качестве индивидуализированных показателей языковой личности. На синтаксическом уровне можно выделить ряд маркеров идиолекта, такие как длина предложений, употребление клишированных конструкций, порядок слов, наличие инверсии, использование вводных слов, типов простых и сложных предложений. Сильным диагностическим маркером является эмоциональный синтаксис, проявляющийся в особом выражении сообщения и оформлении мысли субъектом речи. Эмоциональный синтаксис обусловлен эмоционально-психическими и интеллектуальными параметрами восприятия и интерпретации фрагментов мира.

Прагматический уровень в идиолекте представляет собой высший уровень проявления языковой личности. Он проявляет мотивы, цели, установки субъекта речи. Прагматической функцией речи является воздействие на адресата, и субъект речи может избирать различные стратегии воздействия – от стратегий преодоления коммуникативного и когнитивного диссонанса участников общения, сотрудничества, до стратегий провокационного включения диссонанса, конфликта, дистанцирования как технологий мобилизации и активизации внимания. Стратегиями воздействия могут быть самопрезентация, дискредитация, пропаганда, агитация, манипуляция, реализация которых контролируется коммуникативными стратегиями и определяется целеустановкой субъекта речи. Маркерами идиолекта на прагматическом уровне могут быть языковые средства поддержания интереса адресата к речевому общению или их отсутствие, ориентация на диалоговое взаимодействие или монологичность речи, оценочность высказываний или объективность изложения, аналитичность или информативность, способы актуализации смыслов и многое другое.

При установлении идиолектного профиля в юридической практике происходит сбор речевого материала субъекта с учетом контекста, его прагматических задач. Эти тексты подвергаются лингвистической экспертизе с целью выявления устойчивых паттернов. Для реализации цели изучаемые образцы речи субъекта сравниваются с типовым речевым профилем с учетом территориальных, социальных, возрастных характеристик стандартного профиля. Производится сравнительный анализ спорного текста, авторство которого необходимо установить, и выявленного идиолектного профиля субъекта-потенциального автора спорного текста. На основе совпадений или расхождений делается вывод о вероятном авторстве.

Выводы

Методика многоуровневного анализа идиолектного профиля субъекта речи может использоваться в судебной лингвистике, например, для атрибуции анонимных текстов, установления plagiarisma или выявления фальсификаций, что актуально в практике анализа дипфейков. Проблемой в реализации многоуровневенного анализа идиолектного профиля является индивидуальная вариативность языковых единиц в идиолекте, понимаемом как индивидуальный вариант языка. Разработка методики анализа идиолектного профиля требует решения проблемы определения границ вариативности языковых признаков идиолекта. Вариативностью предполагает изменяемость языковых признаков идиолекта изменяться в зависимости от различных факторов (темы, стиля, модуса и др.), но в определенных количественных пределах, так, чтобы вариативность оставалась внутрииндивидуальной и осуществлялась в границах одного идиолекта, не выходя за его пределы. Актуальность проблемы для юридической практики и теории очень высокая: дипфейки создают новые риски для достоверности доказательств, нарушают права личности, поэтому нужны методы их надежной идентификации, такие как анализ идиолектных маркеров с учётом возможной вариативности языковых единиц в идиолекте. Точность и операционность методики важны для верных правовых решений в условиях цифровизации.

Список источников

1. Бабушкина Е.А. Речевой портрет личности: фонетические характеристики // Вестник Бурятского государственного университета. Язык, литература, культура. 2012. № 11. С. 18 – 24. URL: <https://cyberleninka.ru/article/n/rechevoy-portret-lichnosti-foneticheskie-harakteristiki>.
2. Батурина М.М. Определение авторства текста на основе стилистического анализа с применением синтаксического и морфологического тегирования // Известия Института инженерной физики. 2023. № 4 (70). С. 53 – 57. URL: <https://elibrary.ru/item.asp?id=56000662>.
3. Кащенко А.Э. Проблема демаркации понятий «идиостиль» и «идиолект» // Вестник филологических наук. 2023. Т. 3. № 6. С. 59 – 64. URL: <https://elibrary.ru/item.asp?id=54224866>.

4. Киселев А.С. О необходимости правового регулирования в сфере искусственного интеллекта: дипфейк как угроза национальной безопасности // Вестник Московского государственного областного университета. Серия: Юриспруденция. 2021. № 3. С. 54 – 64. URL: <https://www.elibrary.ru/item.asp?id=47182022>.
5. Лебедев А.А. Анализ последовательностей частей речи и категория идиостиля // Ученые записки Петроводского государственного университета. 2021. № 5. С. 65 – 71. URL: <https://cyberleninka.ru/article/n/analiz-posledovatelnostey-chastey-rechi-i-kategoriya-idiostilya>.
6. Русаяева Е.Ю., Ахобадзе Г.Н. Некоторые особенности литературных текстов при их сопоставлении для определения их авторства // Информационные технологии и вычислительные системы. 2024. № 2. С. 74 – 85. URL: <https://elibrary.ru/item.asp?id=67948808>.
7. Сальников В.Б. Соотношение понятий «идиолект / идиостиль» и «речевой портрет» в современном отечественном языкоznании // Вестник Бурятского государственного университета. Филология. 2025. № 1. С. 21 – 29. URL: <https://www.elibrary.ru/item.asp?id=80492692>.
8. Суэтин В.Ю. Применение частотных характеристик для определения авторства литературных текстов // Вестник Тверского государственного университета. Серия: Прикладная математика. 2022. № 2. С. 84 – 89. URL: <https://elibrary.ru/item.asp?id=49249011>.

References

1. Babushkina E.A. Speech portrait of a person: phonetic characteristics. Bulletin of the Buryat State University. Language, literature, culture. 2012. No. 11. P. 18 – 24. URL: <https://cyberleninka.ru/article/n/rechevoy-portret-luchnosti-foneticheskie-harakteristiki>.
2. Baturin M.M. Determination of text authorship based on stylistic analysis using syntactic and morphological tagging. Bulletin of the Institute of Engineering Physics. 2023. No. 4 (70). P. 53 – 57. URL: <https://elibrary.ru/item.asp?id=56000662>.
3. Kashchenko A.E. The Problem of Demarcation of the Concepts of "Idiostyle" and "Idiolect". Bulletin of Philological Sciences. 2023. Vol. 3. No. 6. P. 59 – 64. URL: <https://elibrary.ru/item.asp?id=54224866>.
4. Kiselev A.S. On the Need for Legal Regulation in the Sphere of Artificial Intelligence: DPfake as a Threat to National Security. Bulletin of Moscow State Regional University. Series: Jurisprudence. 2021. No. 3. P. 54 – 64. URL: <https://www.elibrary.ru/item.asp?id=47182022>.
5. Lebedev A.A. Analysis of Sequences of Parts of Speech and the Category of Idiostyle. Scientific Notes of Petrozavodsk State University. 2021. No. 5. P. 65 – 71. URL: <https://cyberleninka.ru/article/n/analiz-posledovatelnostey-chastey-rechi-i-kategoriya-idiostilya>.
6. Rusyaeva E.Yu., Akhobadze G.N. Some Features of Literary Texts When Comparing Them to Determine Their Authorship. Information Technologies and Computing Systems. 2024. No. 2. P. 74 – 85. URL: <https://elibrary.ru/item.asp?id=67948808>.
7. Salmikov V.B. The Relationship between the Concepts of "Idiolect / Idiostyle" and "Speech Portrait" in Modern Russian Linguistics. Bulletin of the Buryat State University. Philology. 2025. No. 1. P. 21 – 29. URL: <https://www.elibrary.ru/item.asp?id=80492692>.
8. Suetin V.Yu. Application of frequency characteristics to determine the authorship of literary texts. Bulletin of Tver State University. Series: Applied Mathematics. 2022. No. 2. P. 84 – 89. URL: <https://elibrary.ru/item.asp?id=49249011>.

Информация об авторах

Маковиц Г.В., доктор филологических наук, профессор, Уральский юридический институт МВД России,
galinavlad@yandex.ru

© Маковиц Г.В., 2025