

## Идентификационные признаки в судебной экспертизе заимствования текста

**Е. В. Новожилова**

*Российский государственный гуманитарный университет  
Миусская пл., 6, 125997, Москва, Россия. E-mail: elena.v.novozhilova@yandex.ru*

В статье в рамках теории криминалистической идентификации, с опорой на труды известного ученого-криминалиста проф. В. Я. Колдина (1925–2020) рассматриваются признаки, с которыми судебный эксперт работает, обнаруживая текстовые заимствования (копипасту и плагиат).

Предлагается оптимальная классификация идентификационных признаков: шесть групп – посимвольное совпадение текста, опечатки, пунктуационные ошибки, грамматические ошибки, речевые и логические ошибки, иные особенности речи. Автор предлагает использовать термин «эrrативный метод» в экспертизе заимствований. Идентификационные признаки рассматриваются с точки зрения их пригодности для исследования – специфичности, выраженности и устойчивости.

При этом подчеркиваются отличия экспертизы текстовых заимствований от традиционной судебной автороведческой экспертизы: разные идентифицируемые объекты (соответственно текст и человек), разные свойства этих объектов и, следовательно, разные признаки. Автор анализирует классификации идентификационных признаков из важнейших научных и методических пособий по судебному автороведению (на русском языке), сопоставляет эти классификации с приводимой в статье и делает вывод о принципиальных различиях комплексов признаков.

В статье резюмируется, что предложенный комплекс признаков обеспечивает значительную идентификационную избыточность материала – по контрасту и с традиционным автороведением, и вообще с большинством судебных экспертиз. Эта избыточность позволяет эксперту-речеведу сделать надежный вывод о неслучайности совпадения признаков, то есть о тождестве заимствованного текста.

**Ключевые слова:** судебное речеведение, судебное автороведение, криминалистическая идентификация, признак, ошибка.

## Identification Features in Forensic Examination of Text Borrowing

**E. V. Novozhilova**

*Russian State University for the Humanities  
6 Miusskaya Sq., Moscow, Russia. E-mail: elena.v.novozhilova@yandex.ru*

In the article, within the framework of the theory of criminalistic identification, based on works of the well-known criminalist scientist Prof. Valentin Koldin (1925–2020), we consider the features that a forensic expert works with when detecting textual borrowings (copy-paste and plagiarism).

An optimal classification of identification features put into six groups is proposed: text coincidence at character level, misprints, punctuation errors, grammatical errors, speech and logical errors, and other speech features. The author suggests using the term "errative method" in the examination of borrowings and considers identification features in terms of their suitability for research – specificity, manifestation, and stability.

At the same time, the differences between the examination of text borrowings and the traditional forensic authorship examination are emphasized: different identifiable objects (text and person, respectively), different properties of these objects, and, consequently, different features. The author analyzes the classifications of identification features represented in the most important scientific and methodological manuals on forensic authorship examination (in Russian), compares these classifications with the one given in the article, and concludes about the fundamental differences of the feature complexes.

The article concludes that the proposed set of features provides a significant identification redundancy of the material – in contrast to both traditional authorship examination and, in general, to the majority of forensic enquiries. This redundancy allows the expert to make a reliable conclusion about the non-randomness of the coincidence of features, that is, about the identity of the borrowed text.

**Key words:** forensic speech examination, forensic authorship examination, criminalistic identification, feature, error.

В судебно-экспертных задачах на заимствование текста (плагиат, копипаста) решается вопрос, являются ли речевые произведения независимо созданными или же они были скопированы одно с другого. Эти задачи предполагают работу с письменной речью *per se* и, на наш взгляд, практически не требуют исследования внетекстовой реальности – языковой личности автора, его речевых навыков и т. п.

Хотя многие филологи и решают подобные задачи через языковую личность (подробнее см., например: [Абрамкина 2017], [Абрамкина 2019]) – с таким подходом трудно согласиться. Во-первых, здесь человек (автор) не является идентифицируемым объектом, а значит, его личность и свойства не имеют значения. Идентифицируемый объект здесь – текст, который заимствуется, а свойства – это вся полнота его речевого выражения. В рамках экспертизы заимствований должно устанавливаться не тождество лица, а лишь тождество заимствованного текста (из чего с неизбежностью будет следовать тождество его авторского состава): здесь процедура идентификации короче, чем в традиционном автороведении. Во-вторых, языковая личность автора не всегда проявляется на письме достаточно ясно, а в случае коллективного авторства определить идиостиль каждого из писавших документ невозможно – при этом копипасту, то есть тождество текста в ряду документов, по-прежнему легко обнаружить. Такие задачи целесообразно решать через признаки самого текста.

Письменный текст на естественном языке похож на папиллярный узор – в том смысле, что каждый независимо возникший текстовый объект уникален. В криминалистике и судебной экспертизе уникальность – это неповторимая совокупность признаков, называемых идентификационными. Идентификационные признаки – «воспринимаемая непосредственно и представленная в форме сигналов любой физической природы информация о свойствах идентифицируемого объекта, которая может быть использована для отождествления» [Колдин 2003: 26]. Неуникальность же означает, что такой совокупности нет – либо возникло случайное совпадение, либо перед нами один и тот же объект в двух разных проявлениях.

Текст многомерен, и эксперту-речеведу необходимо учитывать его признаки, содержащиеся в разных информационных полях. И важнейшее информационное поле – это сам текст как ограниченная последовательность символов и пробелов. В данном случае, несмотря на то что текст – это знаки и значения, эксперт должен рассматривать последовательность не как знаковую, а как ряд физических сигналов (сигнальная форма выражения идентификационной информации) [Колдин 2003: 28, 34]. Этому информационному полю соответствует важнейший признак – *посимвольное совпадение* текста или фрагмента текста.

Все прочие признаки требуют учета значений.

Промежуточное положение между посимвольным совпадением и остальными признаками занимают *опечатки* (*описки*). Здесь эксперт работает с тем же информационным полем – текстом как последовательностью сигналов, – но для того чтобы обнаружить признак, требуются специальные филологические знания и анализ знаков и значений: необходимо отличить опечатку от правильного написания, с одной стороны, и от ошибки правописания – с другой.

Опечатка – это непреднамеренная ошибка при письме, печатании, компьютерном наборе текста. Ее результат – пропуск буквы, лишняя буква, отсутствие или кратное увеличение пробела между словами, лишний пробел между словом и знаком препинания, пропуск, искажение, удвоение слова и т. п. Однотипные опечатки на протяжении текста свидетельствуют о том, что автор ошибся при использовании функции автозамены. Разновидностью опечатки является так называемая *глазная ошибка*, зрительно похожая на верное написание («гнедой конб»).

Прочие признаки связаны с традиционными для судебного речеведения информационными полями и с морфемным, лексическим и синтаксическим уровнями языка. Поиск здесь сводится к так называемому эрративному подходу, в основе которого – различие нормы и отклонения в речи. Вслед за С. И. Крассой [Красса 2023: 107–108] назовем этот поиск методом эрративного анализа – или можно назвать еще короче, эрративным методом.

Норма, нормативное правописание – это то, к чему пишущий и сознательно, и неосознанно стремится. Такое стремление формируют три обстоятельства:

1. Языковая норма (лингвистическое понятие) – «совокупность стабильных и унифицированных языковых средств и правил их употребления, сознательно фиксируемых и культивируемых обществом» [ЛЭС 1990: 337]. Это и сами языковые средства, и рамки их употребления, и тенденции их отбора и изменения. Языковой норме свойственна историческая устойчивость, благодаря которой создается национальный литературный язык; т. е. это максимально широкое и мощное явление.

2. Добросовестность при общении. Прежде всего это принцип кооперации Г. Грайса – один из его постулатов требует для успешного общения выражаться ясно, избегать непонятных и неоднозначных высказываний [Грайс 1985: 223]. Наряду с этим люди стремятся соблюдать речевой этикет.

3. Обучение правописанию. В частности, каждый носитель языка добровольно (или не вполне добровольно) проходил такое обучение в средней школе – и получил информацию о норме как знания или как речевые автоматизмы.

Однако абсолютная норма представляет собой недостижимый идеал, и любой реальный текст будет отклоняться от нее. Эти отклонения бесконечно разнообразны, и не имеет значения, являются ли они систематическими или несистематическими (errors vs mistakes) [Corder 1967: 166–167]. Их возникновение в речи практически не зависит от воли автора и не регулируется им: в отличие от нормы к отклонениям не стремятся, они почти всегда возникают вопреки желаниям пишущего, поэтому имеют высокую информационную значимость и могут использоваться в качестве идентификационных признаков при решении вопроса о тождестве текста.

Помимо опечаток такими признаками являются:

- пунктуационные ошибки;
- грамматические ошибки;
- речевые ошибки;
- логические ошибки;
- иные особенности речи.

Это простейшая и наиболее удобная классификация, основанная на отраслевом делении теоретической лингвистики. Кроме того, это вариант общепринятой практической классификации ошибок: аналогичные группы выделяют, в частности, учителя русского языка при проверке и оценке заданий на школьном выпускном экзамене – ЕГЭ.

*Пунктуационные ошибки* возникают при неправильном употреблении знаков препинания, используемых в русском языке, – знаков завершения (точка, вопросительный и восклицательный знаки, многоточие), разделения (запятая, двоеточие, точка с запятой, тире) или выделения (запятая, кавычки, скобки, тире). Знаки препинания выполняют в письменной речи две функции. Смыслоразличительная функция помогает верно передать суть, придать однозначный смысл высказыванию: без знака препинания фраза осталась бы непонятной (ср. известное «Казнить нельзя помиловать»). Интонационно-экспрессивная функция – пунктуация в конце предложения обозначает цель высказывания (сообщение, вопрос или побуждение к действию) и интонацию речи, в середине предложения расставляет интонационные и эмоциональные акценты (восхищение, недовольство, радость, удивление и др.). В русском языке употребление знаков препинания в смыслоразличительной функции жестко регламентировано, и отклонения от нормы считаются ошибками, а употребление в интонационно-экспрессивной функции не подчиняется строгим правилам.

*Грамматические ошибки* появляются из-за невладения правилами правописания. Это ошибки в структуре языковой единицы (слова, словосочетания или предложения), то есть нарушения словообразовательных, морфологических или синтаксических норм – например ошибочное словообразование («подчерк»), нарушения норм согласования или управления слов («группа увлекались», «все удивились его красотой»), ошибки в построении сложного предложения («Книга научила меня ценить дружбу, которую я прочитал еще в детстве») и т. п.

*Речевые ошибки* возникают при неправильном использовании лексических ресурсов русского языка (неправильном словоупотреблении), когда автор не полностью понимает то или иное слово, выражение. Нарушения лексических норм – употребление слова в несвойственном ему значении («благодаря пожару лес сгорел»), плеоназм («мертвый труп»), неразличение паронимов («были приняты эффективные меры») и мн. др.

В ту же группу следует поместить и *логические ошибки*, поскольку очень часто такая ошибка одновременно оказывается и речевой (ср. тот же пример «благодаря пожару лес сгорел», многие случаи плеоназма и т. д.). К этой категории относятся подмена тезиса, круговое рассуждение, ложная аналогия, доказательство через то же самое и др.

Последняя группа – *иные особенности речи*, все то, что не является нормой, однако и к ошибкам с точки зрения русского языка их не отнести: например, окказионализмы, фактические ошибки, различные графические особенности, такие как написание отдельных слов КАПСЛОКОМ, и др. Иные особенности речи могут создаваться намеренно, но цель пишущего (за исключением нарочитого искажения слов – «падонкаффского языка») обычно состоит не в том, чтобы отклониться от нормы.

Деление на шесть групп (посимвольное совпадение, опечатки, пунктуационные ошибки, грамматические ошибки, речевые и логические ошибки, иные особенности речи) – это специальная классификация идентификационных признаков письменной речи для нужд экспертизы заимствований. Такая прикладная классификация может и должна отличаться от научной [Колдин 2003: 258, 264] [Вул 1977: 27]. Названные нами признаки являются частными, то есть не относятся к речевому объекту в целом, и не зависят друг от друга: даже если очевидно, что текст написан весьма неграмотным человеком, невозможно предсказать последовательность его ошибок.

Идентификационный признак пригоден для использования в судебной экспертизе, если он специфичен (редок, оригинален), выражен и устойчив.

Для оценки специфичности нам необходимо пересмотреть традиционный лингвистический взгляд на ошибки. Принято считать, что в идиостиле автора систематические ошибки (обусловленные знанием/незнанием языка, уровнем языковой компетенции) закономерны, а несистематические, такие как опечатки и описки, случайны. Однако не в идиостиле в целом, а в каждом конкретном месте конкретного текста крайне трудно выявить непосредственную причину возникновения отклонения. Это связано еще и с тем, что само слово-«субстрат» ошибки появляется в тексте не закономерно. В целом условия формирования и распределения ошибок в единичном объекте нужно признать случайными. Для оценки специфичности здесь нужны количественные методы [Колдин 2003: 175, 274]. И в отсутствие точных данных о встречаемости тех или иных отклонений можно воспользоваться очень упрощенным, но математически корректным рассуждением, что частота слова-«субстрата» с ошибкой всегда меньше, чем частота того же слова без ошибки (ибо при большем значении это уже не ошибка, а узус). Если слова как признаки

высокоспецифичны – значит, ошибки тем более.

Здесь нужно уточнить, что однотипные ошибки (например, «он дерётся» / «пакет рвётся») учитываются в экспертизе заимствований как разные ошибки, а не как одна и та же.

Выраженность ошибок несомненна для профессионального филолога – причем чем грубее ошибка или описка, тем она выраженной. И наоборот, любая нормативная речь оказывается в этом отношении «стертой».

Устойчивость признаков гарантирует материальный носитель. При копировании файлов с электронного на электронный носитель сохраняется тождество и речевого объекта в целом, и его содержимого (текста), но при конвертации или при копировании на неэлектронный носитель они могут видоизмениться и утратить часть признаков. Поэтому стоит обращаться с доказательствами – речевыми объектами так же, как и с вещественными доказательствами. Устойчивость идентификационных признаков письменной речи такова, что редактирование, правка влияет на них незначительно: к утрате приводит только масштабный рерайт. Объект, происхождение которого связано с иным объектом (скопированный текст), даже после правки наследует признаки исходного; объект, представляющий собой частичную переработку, воспроизводит их в той степени, в которой переработка его не затронула.

Еще одно качество признака – его вариационность, то есть способность отражаться в следах по-разному. Здесь экспертам-речеведам повезло, поскольку при копировании текста вручную вариационность признаков минимальна, а при заимствовании способом *copy-paste* она вообще отсутствует.

В традиционном автороведении используются иные перечни признаков. Иначе и быть не может, поскольку здесь другой идентифицируемый объект – человек, свойства которого – языковые и мыслительные навыки. А значит, и признаки, которые ищет эксперт, несут иную информацию.

В 1970-е гг. С. М. Вул, обобщая современные ему наработки криминалистов, предложил деление на общие и частные и, с другой стороны, на признаки языковых и интеллектуальных навыков [Вул 1973: 14–19]. Под общими признаками языковых навыков он понимал степень их развития, «показатель того, в какой мере данное лицо владеет определенными нормами литературного языка», под частными – устойчивые нарушения, предпочтение определенных языковых средств и устойчивое их соотношение в типичных ситуациях общения. Частные признаки составили пять групп: признаки орфографических, пунктуационных, лексико-фразеологических, синтаксических и стилистических навыков. (Позже автор уточнил, что следует объединить в группу признаки не одних только синтаксических, а всех грамматических навыков [Вул 2019: 213].) Признаки интеллектуальных навыков ученому не удалось систематизировать полностью, и он описал только их формально-логическую часть: общие признаки формально-логических навыков – это тоже степень развития, общая способность к выражению мыслей в логической форме, а частные признаки – устойчивые логические ошибки, повторяющиеся нарушения определенных правил логики, предпочтительное использование тех или иных логических элементов и структур, смысловые акцентуации и исправления.

В методическом пособии 2007 г. С. М. Вул более подробно говорит о признаках интеллектуальных навыков письменной речи, выделяя навыки восприятия действительности, переработки информации, навыки акцентуации, аргументации и оценки [Вул 2007: 16–17]. Здесь же он предлагает систематизировать письменно-речевые ошибки, однако без связи с признаками. Все ошибки, по Вулу, делятся на смысловые (фактические и, как их называет автор, «лингвострановедческие», то есть ошибки иностранцев), логические и собственно языковые или речевые. Последнюю категорию можно систематизировать по уровням языка – либо по типу того правила, которое не соблюдается: нарушение системы («так нельзя говорить»), нормы («так не принято говорить по-русски (хотя и можно)») или узуса («так не следует говорить в данной ситуации письменного общения») [Вул 2007: 19].

А. Ю. Комиссаров также различает общие и частные признаки – но и те и другие относит к языковым навыкам и делит частные на пять групп: признаки синтаксических, лексико-фразеологических, стилистических, формально-логических и интеллектуальных навыков [Комиссаров 2000: 23–24].

Методическое пособие ЭКЦ МВД предлагает классификационную рамку, похожую на классификацию Вула, однако вместо языковых и интеллектуальных здесь уже четыре типа признаков: смысловые, текстологические (имеются в виду структурные), языковые и психолингвистические; последние «отражают особенности психики и мышления автора» [Рубцова и др. 2007: 13], а на самом деле касаются субъективно-модальных характеристик речи, некоторых содержательных особенностей и выражения эмоций. Языковые признаки делятся на орфографические, пунктуационные, лексические, фразеологические, стилистические и синтаксические, а в дальнейших практических указаниях продолжается их дробление. Чтобы повысить возможности индивидуализации, эксперты МВД предлагают комплексировать признаки на трех уровнях: набора норм (НН), набора свойств норм (НС) и набора языковых средств выражения свойств норм (НСВ).

В целом методическая мысль традиционного автороведения движется в сторону признаков, недоступных непосредственному наблюдению, и нечетко разграниченных групп признаков.

Экспертиза заимствований отличается от него тем, что в ней:

- учитываются признаки не навыков, а текста;
- используются главным образом частные языковые признаки (общие чаще всего бесполезны);
- неприменима корреляция между признаками как отдельный признак языкового навыка (поскольку не важны навыки);
- как уже говорилось выше, однотипные ошибки не объединяются при подсчете, а считаются отдельно;

- повторяющиеся одинаковые ошибки тоже считаются отдельно, здесь неприменима повторяемость как показатель устойчивости и невариационности признака (устойчивость и вариационность обусловлены иными обстоятельствами).

Итак, встречаемость ошибки как признака – количественно меньше, чем частота того же слова-«субстрата» без ошибки. Данные по словам можно получить из Частотного словаря русского языка [Ляшевская, Шаров 2009] – частоты наиболее употребительных служебных слов не превышают нескольких процентов («и» – 3,6%, «в» – 3,1%), знаменательных еще меньше («быть» – 1,2%, «год» – 0,4%, «новый» – 0,1%), а в среднем частота слова составляет тысячные доли процента. Поскольку общая частота взаимно независимых признаков определяется по теореме умножения вероятностей, то при простом подсчете становится понятно: комбинации ошибок образуют чрезвычайно редкие совокупности, а значит, обладают высочайшей идентификационной значимостью. При решении экспертной задачи на заимствование речевед почти всегда сталкивается с огромной идентификационной избыточностью материала – по контрасту и с традиционным автороведением, и вообще с большинством судебных экспертиз. Эта избыточность позволяет сделать надежный вывод о неслучайности совпадения признаков, то есть о тождестве заимствованного текста.

*Автор выражает благодарность Татьяне Владимировне Бердниковой, к. филол. н., эксперту ФБУ РФЦСЭ при Минюсте России, за помощь в подборе библиографии для этой статьи.*

## Литература

- Абрамкина Е. Е. Автороведческая экспертиза протокола допроса: основные особенности и методика анализа / Вестник Томского государственного университета. - 2017. - № 415. - С. 158-163.
- Абрамкина Е. Е. Идентификационные признаки протокола допроса и методика автороведческого анализа / Вестник Новосибирского государственного университета. Серия Лингвистика и межкультурная коммуникация. - 2019. - Т. 17. - № 3. - С. 97-108.
- Вул С. М. Криминалистическое исследование признаков письменной речи. Киев, 1973.
- Вул С. М. О классификации идентификационных признаков письменной речи в судебно-автороведческой экспертизе / Вестник Университета им. О. Е. Кутафина (МГЮА). - 2019. - № 5. - С. 213-217.
- Вул С. М. Судебно-автороведческая идентификационная экспертиза: Методические основы. Харьков, 2007.
- Вул С. М. Теоретические и методические вопросы криминалистического исследования признаков письменной речи. М., 1977.
- Грайс Г. П. Логика и речевое общение / Новое в зарубежной лингвистике. Вып. 16. Лингвистическая прагматика. М., 1985. С. 217-237.
- Колдин В. Я. Судебная идентификация. М., 2003.
- Комиссаров А. Ю. Криминалистическое исследование письменной речи. М., 2000.
- Красса С. И. Методика определения меры сходства между объектами речеведческих экспертиз / Юрислингвистика. - 2023. - № 27 (38). - С. 106-110.
- Лингвистический энциклопедический словарь (ЛЭС) / Гл. ред. В. Н. Ярцева. М., 1990.
- Ляшевская О. Н., Шаров С. А. Новый частотный словарь русской лексики. 2009. URL: <http://dict.ruslang.ru/freq.php?>
- Рубцова И. И., Ермолова Е. И., Безрукова А. И. и др. Комплексная методика производства автороведческих экспертиз: Методические рекомендации. М., 2007.
- Corder S. P. The significance of learner's errors / International Review of Applied Linguistics in Language Teaching. - 1967. - Т. 5 (4). - С. 161-170.

## References

- Abramkina, E. E. (2017). The authorship examination of the interrogation protocol: basic special aspects and methods of analysis. Bulletin of Tomsk State University, 415, 158-163 (in Russian).
- Abramkina, E. E. (2019). Identification features of the interrogation protocol and the methodology of the authorship analysis. Bulletin of Novosibirsk State University, Linguistics and Intercultural Communication Series, 17, 3, 97-108 (in Russian).
- Corder, S. P. (1967). The significance of learner's errors. International Review of Applied Linguistics in Language Teaching, 5, 4, 161-170 (in English).
- Grice, H. P. (1985). Logic and conversation. New in foreign linguistics, 16, 217-237 (in Russian).
- Koldin, V. Ya. (2003). Judicial identification. Moscow (in Russian).
- Komissarov, A. Yu. (2000). Forensic examination of written speech. Moscow (in Russian).
- Krassa, S. I. (2023). Methodology for determining the measure of similarity between the objects of speech examinations. Legal Linguistics, 27(38), 106-110 (in Russian).
- Linguistic Encyclopedic Dictionary (1990) / Ed. by V. N. Yartseva. Moscow (in Russian).
- Lyashevskaya, O. N., Sharov, S. A. (2009). New frequency dictionary of Russian vocabulary. Available from: <http://dict.ruslang.ru/freq.php?> (in Russian).

- 
- Rubtsova, I. I., Ermolova, E. I., Bezrukova, A.I. et al. (2007). A comprehensive methodology for the production of forensic authorship examinations: Guidelines. Moscow (in Russian).
- Vul, S. M. (1973). Forensic examination of features of written speech. Kiev (in Russian).
- Vul, S. M. (1977). Theoretical and methodological issues of forensic examination of features of written speech. Moscow (in Russian).
- Vul, S. M. (2007). Forensic authorship identification expertise: Methodological foundations. Kharkov (in Russian).
- Vul, S. M. (2019). On the classification of the identification features of written speech in the forensic authorship examination. Bulletin of Kutafin Moscow State Law University (MSAL), 5, 213-217 (in Russian).
- 

**Citation:**

Новожилова Е.В. Идентификационные признаки в судебной экспертизе заимствования текста // Юрислингвистика. – 2023. – 30. – С. 83-88.

Novozhilova E.V. (2023) Identification Features in Forensic Examination of Text Borrowing. Legal Linguistics, 30, 83–88.



This work is licensed under a Creative Commons Attribution 4.0. License

---