

Методика определения меры сходства между объектами речеведческих экспертиз

С. И. Красса

*Ассоциация «Центр независимой экспертизы "Спектр"»
Ставрополь, Россия. E-mail: skrassa@mail.ru*

В статье описываются исследовательские процедуры с целью определения меры сходства между объектами речеведческих экспертиз. Предлагаемая методика основывается на сочетании качественных и количественных методов, а также применении компьютерного инструментария. Основной задачей данного исследования является описание способов получения максимально надежных результатов, которые могут быть проверены другим специалистом. Для проведения исследования выделяются три множества: исходный объект (спорный текст), объект, сопоставляемый с исходным, и объект-конструкт, включающий в себя характеристики, общие для исходного и сопоставляемого множеств. В качестве объектов могут быть протоколы допроса, тексты, проверяемые на наличие плагиата, наименования товарных знаков, газетные публикации, другие речевые продукты. Качественные процедуры предполагают выделение характеристик, описывающих объекты, тогда как количественные процедуры позволяют проводить нормализацию параметров, выявленных в объектах. Выделение характеристик проводится экспертным путем, с опорой на принятые в лингвистике классификации. Анализ текстов с идентичным содержанием проводится путем выделения тождественных фрагментов и подсчета их объемов. Анализ сходства словесных наименований проводится на базе фонетических, графических, семантических и ассоциативных параметров. Анализ текстов на возможное авторство проводится с помощью лексических, морфологических и синтаксических данных. На основе полученных численных показателей проводится вычисление коэффициентов сходства. Характеристики, которые предполагают ответ «да/нет», маркируются единицей при положительном ответе и нулем при отрицательном. Характеристики, имеющие числовое значение, сопоставляются с использованием коэффициентов корреляции и признаются идентичными при значении 0,7 и выше. Проведение цифровых аналитических процедур по возможности основывается на использовании компьютерных сервисов. Полученные числовые данные представляют собой коэффициенты Жаккара, Серенсена, Кульчинского и Отиаи, среди которых первый всегда ниже трех последующих, демонстрирующих относительную стабильность.

Ключевые слова: лингвистическая экспертиза, автороведческая экспертиза, сходство текстов, тождество текстов.

The Procedure of Determining Degree of Similarity between Objects in Linguistic Examinations

S. I. Krassa

*Association "The Independent Expertise Center " Spektr "
Stavropol, Russia. E-mail: skrassa@mail.ru*

The article describes the research procedures for determining the degree of similarity between the objects of linguistic examination. The procedure is based on the combination of qualitative and quantitative methods and the use of computer tools. The main objective of this study is to describe ways of obtaining the most reliable results, which can be verified. For carrying out of research three sets are allocated: initial object, object compared with initial, and object-construct, which includes the characteristics common for the compared sets. The objects may be interrogation protocols, texts checked for plagiarism, trademark names, newspaper publications and other. Qualitative procedures imply singling out characteristics describing objects, while quantitative procedures allow normalization of parameters identified in the objects. The extraction of characteristics is carried out by expertise, relying on the classifications accepted in linguistics. The analysis of texts with identical content is carried out by selecting identical fragments and estimating their volumes. The analysis of the similarity of trademarks is carried out on the basis of phonetic, graphic, semantic and associative parameters. The analysis of texts for

possible authorship is carried out with the help of lexical, morphological and syntactic data. On the basis of the obtained numerical indicators, similarity coefficients are calculated. Characteristics that involve a selection of a yes/no answer are marked with 1 for a positive answer and zero for a negative one. Characteristics that have a numerical expression are compared using correlation coefficients and are recognized as identical when the value is 0.7 or higher. Numeric analytical procedures are generally based on the use of computer services. The numeric data obtained are represented by the coefficients of Jaccard, Sørensen, Kulczynski and Ochiai.

Key words: linguistic examination, authorship examination, similarity of texts, identity of texts.

Рассматривая отношения гуманитарных и естественных наук, N. Christakis отмечал: гуманитарии не поступают так, как это делают представители естественных наук, заявляя об исследовании области до «разумного уровня определенности», чтобы затем перейти к иным рубежам [Christakis 2015]. На наш взгляд, «уровнем определенности» для некоторых речеведческих экспертиз является, в частности, установление сходства между языковыми объектами на основе сочетания воспроизводимых качественных и количественных методов.

В судебно-экспертной практике нередко встречается анализ протоколов допроса, «значительная часть которых выполнена “под копирку” и имеет аналогичные ошибки. Такие экспертные задачи смыкаются в определенной мере с автороведческой экспертизой, в том смысле, что выясняется, кто автор текста – свидетель (потерпевший, подозреваемый) или следователь (дознатель) – лицо, составлявшее протокол» [Красса 2018: 82]. В подобных исследованиях обычно ставятся задачи выявления в текстах идентичных (тождественных) фрагментов, а также аналогичных ошибок. Сходной задачей является установление процента идентичности судебного решения (приговора) и обвинительного заключения, а также иных экспертных действий с целью демонстрации текстовых заимствований.

Для решения указанных задач используется метод сопоставительного анализа в его разновидности корректорской считки (сличения) [Новожилова 2021], а также статистический анализ. Для ответа на второй вопрос применяется метод эрративного анализа (выявления ошибок и их интерпретации). Идентичные (тождественные) фрагменты имеют такое же лексическое содержание и грамматические формы, написание и пунктуацию и, следовательно, одинаковое смысловое содержание [Красса 2019: 82]. В данном случае мы уходим от разграничения «полностью совпадающие / частично совпадающие фрагменты» [Бойцов, Кулешова, Лизоркин 2018].

Проведенное сопоставление (сличение) текстов (например, протоколов опроса, допроса и т. п.) оформляется, как правило, в таблицах. Совпадающие части текстов разбиваются на фрагменты и нумеруются. Затем проводится подсчет объема фрагментов с помощью стандартной офисной программы Microsoft Word, сервис «Статистика» (количество слов, знаков без пробелов). Различающиеся части анализируемых фрагментов выделяются для наглядности фоном.

Частично эту работу можно возложить на компьютерные сервисы, например Copyscape [<https://www.copyscape.com/compare.php>], Embedica [<https://compare.embedika.ru>], Draftable [<https://draftable.com/compare/>], Панда-Копирайтинг [<https://petr-panda.ru/sravnitiv-teksty/>], Copyleaks [<https://app.copyleaks.com/text-compare>] (см. обзор этих программ в выступлении П. Е. Беловой на Всероссийской научно-практической конференции «Юридическая лингвистика: языковые аспекты права», Барнаул, 14 октября 2022 г.).

Использование указанных программ помогает специалисту в выделении идентичных фрагментов сопоставляемых текстов, а также подсчете слов и знаков. Вместе с тем это не снимает ответственности с эксперта за результаты исследования и их интерпретацию (см. о проблеме «черного ящика» в судебных делах по автороведению [Красса 2020]). В тех случаях, когда программы открывают свои алгоритмы, также остаются вопросы. Например, в одной из программ «расчет числового параметра степени совпадения сравниваемых абзацев производится на основании вычисления расстояния Левенштейна» [Бойцов, Кулешова, Лизоркин 2018: 49], что, на наш взгляд, математически вполне оправданно. В то же время в итоговую формулу расчета процента совпадения включены по два коэффициента (в числитель и знаменатель). Эти коэффициенты «подобраны эмпирически (на основании репрезентативных выборок из результатов сравнительных анализов текстов небольших объемов по ранее проведенным экспертным исследованиям)» [Бойцов, Кулешова, Лизоркин 2018: 49]. Такой подход априори ставит вопросы относительно объективности названных коэффициентов, прозрачности их получения и возможности их экстраполяции на иные тексты.

В результате анализа текстов (вручную или с помощью компьютерных программ) в распоряжении эксперта оказываются 3 текста: текст А – исходный в том или ином отношении (например, более ранний по дате создания), текст В, который сопоставляется с исходным, и текст С, содержащий фрагменты, общие с текстами А и В. Далее проводится подсчет объема каждого их текстов. Например (данные взяты из реального исследования), если объем текста А – 26743 знака без пробелов, объем текста В – 15104 знаков, текста С – 12907 знаков, то эксперт отвечает примерно таким образом: «В текстах А и В имеются идентичные (тождественные) фрагменты, доля которых составляет 56 % в тексте А и 85 % в тексте В». Возможна также формулировка: «56 % из текста А входит в текст В и составляет в последнем около 85 % от его объема». Доля общего текста разная из-за различия в объемах текстов, представленных на исследование.

Чтобы избежать двойственности обозначения общей доли в исследуемых текстах, следует использовать коэффициенты сходства. Сходство – «отношение, родственное равенству. При наличии у пары объектов хотя бы

одного общего признака можно говорить о С. объектов этой пары. Ввиду многообразия признаков на одной паре могут индуцироваться разные отношения С., а ввиду повторяемости признаков – одно отношение С. на разных парах [Сходство 1983: 666].

Для выявления меры сходства могут использоваться коэффициенты, например коэффициент Жаккара. Это первый известный коэффициент сходства, который в различных модификациях и записях активно используется для решения статистических задач. Формула данного коэффициента $K_{ж} = C/(A+B-C)$. Для приведенного выше примера значение коэффициента равно 0,45. Однако это не единственный коэффициент сходства, причем его значения, как позывает наша практика, ниже других аналогичных коэффициентов: так, в данном кейсе коэффициент Серенсена $K_{с} = 2C/(A+B) = 0,62$; коэффициент Кульчинского $K_{к} = C/2*(1/A+1/B) = 0,67$; коэффициент Отиаи $K_{о} = C/\sqrt{AB} = 0,64$.

При других показателях объема (размера) текстов коэффициенты (кроме коэффициента Жаккара) демонстрируют более близкие друг к другу значения. Например, при показателях в количестве слов $A = 813$, $B = 833$, $C = 823$ коэффициенты следующие: $K_{с} = 0,84$, $K_{к} = 0,85$, $K_{о} = 0,85$. Вероятно, стабильность названных выше трех коэффициентов зависит от соотношения объемов исследуемых текстов (эти объемы не должны значительно отличаться друг от друга), а также от размера доли общего текста в них.

Эрративный анализ показывает аналогичные ошибки в исследуемых текстах. При всем несистемном характере этого явления, указание на одни и те же ошибки представляет собой весьма весомый аргумент тождественности фрагментов исследуемых текстов. Часто среди таких ошибок имеет место отсутствие необходимых знаков препинания между частями сложного предложения, внутри осложненного предложения, грамматические ошибки в употреблении падежных форм (в частности, с предлогами *благодаря*, *согласно*, *вопреки*), отдельные орфографические и иные ошибки и опiski.

Следующим кейсом установления меры сходства между объектами является сопоставительное исследование наименований товарных знаков. Методика данного исследования описана в [Новичихина, Стернин 2013], ее недостатки и пути их преодоления представлены в [Радбиль, Красса 2022]. Анализ сходства словесных наименований А и В проводится на основе исследования фонетических, графических, семантических и ассоциативных параметров.

В фонетическом анализе вначале приводится фонетическая транскрипция исследуемых слов, затем проводится сопоставление каждого из звуков на основе фонетико-фонологических характеристик. Названные характеристики могут быть не только артикуляционными: ряд, подъем, лабиализованность (а также ударность, если нет отдельного супrasegmentного анализа) для гласных, способ, место, сонорность, палатализованность для согласных, но и акустическими. Причем акустические характеристики представляются более удобными в данном случае, во-первых, в силу их универсальности (для гласных и согласных), во-вторых, в силу их дихотомичности.

Если в предыдущем кейсе числовые значения имеют естественный, не условный характер (объем текста измеряем тем или иным способом), то в данном случае необходима нормализация качественных параметров, которая проводится следующим образом. Звукам наименования А (см. [Радбиль, Красса 2022]) присваивается значение единицы (1) по каждому из параметров; звукам наименования В также 1 при совпадении характеристик и 0 при их различии. Например, звук [л] в одинаковой позиции в слове А и В (ALDBIO и ALBEA) описывается как вокальный, неконсонантный, низкий, компактный, недиезный, небемольный, непрерывный, нерезкий, звонкий, и каждой характеристике присваивается по 1 для каждого слова. Звуки [л] и [и] совпадают в характеристиках «вокальный, неконсонантный, небемольный, непрерывный, нерезкий, звонкий» и получают по 1 баллу. Характеристики «высокий, компактный, недиезный» отличают [и], поэтому данный звук за них ничего не получает (или получает по нулю), тогда как звук [л] за корреляты названных характеристик (низкий, диффузный, диезный) получает по единице. Множество А формируется путем сложения баллов (единиц), которыми положительно маркируются звуки слова А; множество В формируется путем сложения баллов (единиц), которыми положительно (так же, как и слова А) маркируются звуки слова В в сравнении со словом А.

Для выявления сходства необходимо еще одно множество – С, оно формируется путем сложения баллов, которыми положительно маркируются звуки слова В в сравнении со словом А – также по одному баллу за признак. Так, для исследуемых наименований со значениями $A = 45$, $B = 44$, $C = 35$ получаем следующие коэффициенты: $K_{ж} = 0,65$, $K_{с} = 0,79$, $K_{к} = 0,79$, $K_{о} = 0,79$. Как и в предыдущем кейсе, коэффициент Жаккара ниже, остальные коэффициенты стабильны или одинаковые. На них и следует опираться эксперту при формулировании выводов. В интерпретации данных квантитативного анализа рекомендуется исходить из значений корреляции, принятых в статистике: $0,2 < r \leq 0,5$ слабая корреляция; $0,5 < r \leq 0,7$ средняя корреляция; $0,7 < r \leq 0,9$ сильная корреляция; $0,9 < r \leq 1$ очень сильная корреляция. Следовательно, коэффициент 0,79 следует рассматривать в качестве индикатора сильного сходства.

Интересно, что фонетический артикуляционный анализ для пары тех же слов дает иные значения коэффициентов: $K_{ж} = 0,5$; $K_{с} = 0,67$; $K_{к} = 0,68$; $K_{о} = 0,67$. По этим данным степень сходства средняя. Различные подходы к анализу фонетических параметров слова открывают возможность исследования меры близости слов иными метриками.

Следующий вид анализа – графический, который также имеет два аспекта. Первый – в определенной мере общие характеристики, не связанные со значением отдельных графических знаков. Это такие характеристики, как тип письма (фонография/идеография), алфавит (кириллица/латиница), графема (буква/диграф), прописная или строчная, шрифт (курсив, полужирный, буквица). Второй аспект связан со значением конкретной графемы и с парным сопоставлением каждой из них: гласная или согласная графема, обозначает звук (не обозначает звука или два звука, имеет иные графические функции – например, обозначение мягкости предшествующего согласного), какую именно фонему обозначает. Для описываемых слов получаются следующие значения коэффициентов $K_{ж} = 0,77$; $K_{с} = 0,87$; $K_{к} = 0,87$; $K_{о} = 0,87$. По этим данным степень графического сходства высокая.

Семантическая параметризация фантазийных наименований представляет определенные трудности ввиду непрозрачности их семантики в отличие от значения узуальных слов. В данном случае было предложено два блока «Морфемика/словообразование» (членимость, производность, морфемное (словообразовательное) строение) и «Лексика» (значение, компоненты семантики, этимология). По результатам анализа указанных выше фантазийных наименований получаем следующие коэффициенты: $K_j = 0,5$; $K_c = 0,67$; $K_k = 0,75$; $K_o = 0,71$. По приведенным данным степень сходства средняя.

И, наконец, завершает линейку сопоставления двух наименований фоносемантический анализ. Данный вид анализа часто вызывает скепсис в филологических кругах. Очевидно, это связано с позицией по отношению к фоносемантике в целом. Тем не менее полагаем, что в случае с неявной, затемненной, неоднозначной семантикой наименований фоносемантические ассоциации важны. Фоносемантический анализ представляет собой валидный, научно обоснованный метод исследования одного из аспектов лингвистической экспертизы товарного знака.

В работе [Новичихина, Стернин 2013] имеет место описание такого анализа применительно к названиям HELMAR и HELMAN. Однако авторы не указали компьютерную программу, с помощью которой они выполняли этот анализ, алгоритм анализа в программе. Судя по представленным результатам, это программа Diatone, монополярная шкала, алгоритм А. П. Журавлева.

Коэффициент фоносемантического сходства (формулировка авторов пособия) определяется следующим образом. Для названий с тождественным количеством фоносемантических признаков он определяется как отношение числа совпадающих фоносемантических признаков к общему числу признаков, выраженное в процентах. Для единиц, имеющих нетождественное количество фоносемантических признаков, коэффициент фоносемантического сходства определяется как удвоенное отношение числа совпадающих фоносемантических признаков к общему числу признаков в обоих сопоставляемых элементах, выраженное в процентах. В этих формулировках вызывает вопросы, во-первых, «тождественное количество фоносемантических признаков»: одинаковое число признаков или число одинаковых признаков? Если у сопоставляемых слов, как отмечается в пособии, по 16 тождественных фоносемантических признаков, то коэффициент фоносемантического сходства будет единица, или 100 %. По мнению авторов пособия, КФС для HELMAR и HELMAN будет $32/34 = 94 \%$, то есть они используют коэффициент Серенсена.

Между тем реальная картина признаков данных слов несколько иная. Признаки МЕДЛЕННЫЙ, КОРОТКИЙ выделяются у слова HELMAN, однако не указываются для слова HELMAR, хотя, по данным той же программы, присутствуют в его фоносемантической структуре практически с теми же показателями, что и в слове HELMAN. Кроме того, и в том и в другом слове есть характеристика МЕДЛИТЕЛЬНЫЙ, не отраженная в анализе в данном пособии. В таком случае все коэффициенты будут равны единице.

Чтобы избежать отмеченных недостатков, необходимо, указав программу, тип шкалы и алгоритм, выделить пары идентичных признаков. Так, для пары ALBEA/ALDBIO это ХОРОШИЙ, СВЕТЛЫЙ, ЛЕГКИЙ, ВЕСЕЛЫЙ, БЕЗОПАСНЫЙ, ЯРКИЙ, ОКРУГЛЫЙ, РАДОСТНЫЙ. ГРОМКИЙ, ДОБРЫЙ отмечены только для первого слова, тогда как ПРОСТОЙ, ГЛАДКИЙ – для второго. Корреляция числовых показателей для пар ХОРОШИЙ ... ДОБРЫЙ колеблется в пределах 0,97 – 0,99, то есть это идентичные показатели. Следовательно, для данной пары наименований $A = 10$, $B = 10$, $C = 8$. В таком случае $K_j = 0,66$; $K_c = 0,8$; $K_k = 0,8$; $K_o = 0,8$. По приведенным данным степень сходства высокая.

Таким образом, с максимально достижимой в современной науке степенью точности и объективности можно утверждать, что в характеризуемой паре наименований высокая степень фонетического (акустического), графического и фоносемантического (ассоциативного) сходства и средняя степень фонетического (артикуляционного) и семантического сходства.

Наконец, установление меры сходства между объектами может быть проведено в рамках экспертизы, которая классификационно относится к автороведческой, однако мы не видим сколь-нибудь значимых ее отличий от лингвистической – по крайней мере в характере анализа, результаты которого приводятся ниже. В качестве объектов рассматриваются два текста из одного номера газеты: один подается как написанный от имени газеты, тогда как второй – письмо читателя в газету. Проблема в двух текстах поднимается одна и та же. Объем первого текста составляет примерно 0,78–0,8 от объема второго текста.

Выделяются лексический, морфологический и синтаксический блоки признаков. В лексический входят длина слова (в знаках), длина слова (в слогах), индекс лексического разнообразия. Затем проводится сопоставление признаков. Например, если длина слова в слогах в тексте А – 2,89, а в тексте В – 2,81, то корреляция этих показателей 0,97 и они рассматриваются как очень близкие; при этом по одному баллу присваивается тексту А, тексту В и тексту С – конструктору, объединяющему в себе общие характеристики А и В.

В морфологический блок вводятся коэффициенты предметности (отношение суммы существительных и местоимений к сумме прилагательных и глаголов), качественности (отношение суммы прилагательных и наречий к сумме глаголов и существительных), активности (отношение суммы глаголов и глагольных форм к количеству слов в тексте), динамизма (отношение суммы глаголов и глагольных форм к сумме существительных, прилагательных и местоимений) [Радбиль, Маркина 2019]. Как представляется, данные коэффициенты довольно полно покрывают морфологическую картину текста, в том числе в аспекте соотношения морфологических категорий, характеризующих стиль автора.

В синтаксический блок входят длина предложения, коэффициент связности текста (отношение суммы предлогов и союзов к числу предложений), число предложений с гипотаксисом и с паратаксисом, индекс туманности Ганнинга (уровень удобочитаемости текста).

В нашем случае были получены следующие результаты: коэффициент Жаккара 0,92, коэффициент Серенсена 0,96, коэффициент Кульчинского 0,96, коэффициент Отиаи 0,96. Специалисты по автороведению полагают, что при значениях коэффициентов больше 0,87 тексты следует рассматривать как принадлежащие одному автору.

Таким образом, установление меры сходства между объектами в речеведческих экспертизах проводится на основе сочетания качественных и количественных методов. Для этого выделяются исходный объект, объект, сопоставляемый с исходным, и объект-конструкт, включающий в себя характеристики, общие для исходного и сопоставляемого объектов. Рассматриваются протоколы допроса, наименования товарных знаков, газетные тексты. Качественные процедуры предполагают выделение характеристик, адекватно описывающих объект, тогда как количественные процедуры позволяют проводить нормализацию данных характеристик в объектах. На основе полученных численных показателей проводится вычисление коэффициентов сходства.

Литература

- Christakis N. A.* Let's Shake Up the Social Sciences / *New York Times. Sunday Review.* July 19, 2013. URL: <http://www.nytimes.com/2013/07/21/opinion/sunday/lets-shake-up-the-social-sciences.html>.
- Бойцов А. А., Кулешова А. А., Лизоркин А. М.* Новый подход к сравнительному анализу текстов в рамках комплексной компьютерно-технической и лингвистической экспертизы / *Теория и практика судебной экспертизы.* – 2018. Том 13. № 3. – С. 47-52.
- Красса С. И.* Первый круглый стол по практикам и стандартам судебного автороведческого анализа (обзор 1) / *Политическая лингвистика.* – 2020. – № 6 (84). – С. 174-187.
- Красса С. И.* Рамки судебной лингвистической экспертизы / *Язык. Текст. Дискурс: Научный альманах Ставропольского отделения РАЛК / Под ред. проф. Г.Н. Манаенко.* Выпуск 17. – Ставрополь, 2018. – С. 49-88.
- Новичихина М. Е., Стернин И. А.* Экспертиза товарного знака: учебное пособие. Воронеж, 2013.
- Новожилова Е. В.* Корректорская считка (сличение) как метод судебного речеведения / *Теория и практика судебной экспертизы в современных условиях: материалы VIII научно-практ. конф. Москва, 28-29 января 2021 г.* – С. 233-235.
- Радбиль Т. Б., Красса С. И.* Методика лингвоэкспертного сопоставительного исследования фантазийных наименований товарных знаков: новые решения / *Национальные и международные тенденции и перспективы развития судебной экспертизы: сборник докладов Научно-практической конференции с международным участием, г. Нижний Новгород, 19-20 мая 2022 г.* – С. 210-218.
- Радбиль Т. Б., Маркина М. В.* Вероятностно-статистические модели в производстве автороведческой экспертизы русскоязычных текстов / *Политическая лингвистика.* – 2019. – № 2 (74). – С. 156-166.
- Сходство / Философский энциклопедический словарь.* М. Гл. редакция: Л. Ф. Ильичёв, П. Н. Федосеев, С. М. Ковалёв, В. Г. Панов. 1983.

References

- Boytssov, A. A., Kuleshova, A. A., Lizorkin, A. M. (2018) New approach to the comparative analysis of texts in the framework of complex computer-technical and linguistic examination. Theory and practice of forensic examination, 13, 3, 47-52 (in Russian).
- Krassa, S. I. (2018). The framework of forensic linguistic examination. Language. Text. Discourse. Scientific almanac of Stavropol branch of RALC, 17, 49-88, Stavropol (in Russian).
- Krassa, S. I. (2020). First round table on the practices and standards of forensic authorship analysis (review 1). Political Linguistics, 6 (84), 174-187 (in Russian).
- Novichikhina, M. E., Sternin, I. A. (2013). Examination of a trademark. Textbook, Voronezh (in Russian).
- Novozhilova, E. V. (2021). Corrective reading (comparison) as a method of forensic speech science. Theory and practice of forensic examination in modern conditions. Materials of VIII scientific-practical conf. Moscow, 28-29 January 2021, 233-235 (in Russian).
- Radbil, T. B., Krassa, S. I. (2022). Technique of linguistic-expert comparative study of fantasy trademark names: new solutions. National and international tendencies and prospects of forensic examination. Collection of reports of scientific-practical conference with international participation, Nizhny Novgorod, 19-20 May 2022, 210-218 (in Russian).
- Radbil, T. B., Markina, M. V. (2019). Probabilistic-statistical models in the production of author's attribution of Russian-language texts. Political Linguistics, 2 (74), 156-166. DOI 10.26170/pl19-02-18 (in Russian).
- Similarity. (1983). Philosophical Encyclopedic Dictionary. Moscow, Editors-in-chief: L. F. Ilyichev, P. N. Fedoseev, S. M. Kovalev, V. G. Panov (in Russian).

Citation:

Красса С. И. Методика установления меры сходства между объектами речеведческих экспертиз // Юрислингвистика. – 2023. – 27. – С. 106-110.

Krassa S. I. (2023). The Procedure of Determining Degree of Similarity between Objects in Linguistic Examinations. Legal Linguistics, 27, 106-110.



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