ИЗДАТЕЛЬСТВО ГРАМОТА

Филологические науки. Вопросы теории и практики
Philology. Theory & Practice

2022. Том 15. Выпуск 1. С. 203-208 | 2022. Volume 15. Issue 1. P. 203-208

ISSN 1997-2911 (print)

Материалы журнала доступны на сайте (articles and issues available at): philology-journal.ru

RU

Интерпретация художественного произведения: корпусный подход

Горожанов А. И., Степанова Д. В.

Аннотация. Исследование ставит перед собой цель определить оптимальный с точки зрения достоверности метод интерпретации художественных произведений в рамках корпусного подхода. Научная новизна заключается в том, что формулируются основные положения универсального метода, который позволит извлекать данные для моделирования пространственно-временной и качественной структур произведения, установить черты идиостиля автора. Полученные результаты показали, что работа с неаннотированными лингвистическими корпусами может быть эффективной при использовании современных программных инструментов обработки естественного языка.

EN

Work of Fiction Interpretation: Corpus Approach

Gorozhanov A. I., Stepanova D. V.

Abstract. The research aims to determine an optimal method for interpreting works of fiction in terms of reliability within the framework of the corpus approach. The scientific originality lies in the fact that the main provisions of the universal method are formulated, which allows extracting data for modelling the spatial-temporal and qualitative structures of a work, and ascertaining the features of the author's idiostyle. The results obtained have shown that working with unannotated linguistic corpora can be effective when using modern software tools for natural language processing.

Введение

Современные компьютерные технологии позволяют накапливать и оперативно обрабатывать сверхбольшие объемы данных, в том числе и для нужд компьютерной лингвистики. Национальные лингвистические корпусы характеризуются как корпусы третьего поколения, или «гигакорпусы», размещенные в Интернете с целью решения проблемы преодоления несовместимости с программным обеспечением пользователя и достижения высокой скорости исполнения поисковых запросов (Солнышкина, Гатиятуллина, 2020, с. 150). Но даже такие вычислительные мощности не позволяют в полной мере решать актуальные задачи корпусной лингвистики, в частности связанные с интерпретацией художественных произведений. А. Н. Баранов, Д. О. Добровольский и Н. А. Фатеева (2021) отмечают, что «использование современных корпусных технологий обработки данных дает возможность более строго формировать модели идиостиля [писателя], усиливая их объективность в результате применения количественных методов» (с. 374), а «современные информационные технологии позволяют радикально модернизировать существующие модели идиостиля», однако в отношении исследования нарративных структур и системы интертекстуальных связей пока что удается достичь меньших результатов, чем в отношении описания лексической и синтаксической характеристик текста (с. 386). Тем более сложным является декодирование идиостиля автора в поэтическом произведении, для которого в современных исследованиях применяется «метод целостного анализа стихотворного произведения, предполагающий применение элементов герменевтики» (Яновская, Чижикова, Золотых, 2019, с. 253). Высокий интерес к проблемам интерпретации художественных произведений, описания идиостиля автора точными методами обуславливает актуальность настоящего исследования.

Планируя проведение корпусных исследований с целью интерпретации художественных произведений (прозы), научному коллективу необходимо определиться с количественной характеристикой привлекаемого или создаваемого корпуса. Это может быть либо национальный («гигакорпус»), либо сбалансированный (специальный) лингвистический корпус.

Другой, по сути ключевой проблемой корпусных исследований является выбор способа разметки, или аннотирования, текстовых данных. По сути, «корпус – это практически любая совокупность языковых фактов, необязательно связанная с компьютерно-опосредованным представлением языка» (Баркович, 2016, с. 8-9).

Научная статья (original research article) | https://doi.org/10.30853/phil20220020

В этой связи корпусами могут считаться также и неаннотированные электронные текстовые массивы, хотя, как правило, лингвистические корпусы являются размеченными (Козлова, 2013, с. 87). Важность принятия решения о выборе аннотированного или неаннотированного корпуса на начальном этапе исследования заключается в следующем. При выборе национального корпуса, который отражает состояние языка как некое среднее на текущий момент времени, применительно к исследованию произведений конкретного писателя, а тем более – к попыткам описания его идиостиля весьма высока вероятность риска ошибочной интерпретации, поскольку «идиостиль рассматривается в коммуникативно-когнитивном аспекте комплексно: как многоплановое проявление мировидения автора в структуре, семантике и прагматике текста» (Болотнова, Болотнов, 2012, с. 188). Оптимальным решением в данном случае является самостоятельное составление сбалансированного корпуса исследуемых текстов, однако здесь, в свою очередь, необходимо точно выбрать характер разметки, поскольку от этого будет зависеть специфика программного обеспечения для исполнения поисковых запросов (Gorozhanov, Guseynova, 2020a, с. 2038). Принятие решения в пользу неаннотированного корпуса, с одной стороны, снимает вопрос о выборе способа тегирования и его проведении, а с другой стороны, ставит не менее острый вопрос о программном инструменте, с помощью которого будут получены необходимые данные из неразмеченного корпуса. Наконец, во всех указанных случаях необходимо учитывать специфику исследуемого языка (Bolshina, Loukachevitch, 2020; Kim, Kwon, 2021).

Сообразно этому формулируются задачи работы:

1. Определить параметры интерпретации художественных произведений.
2. Определить тип лингвистического корпуса, необходимого для исследования всех заданных параметров.
3. Произвести отбор подходящих программных решений.
4. Провести предварительную апробацию метода.

Практическая значимость исследования заключается в том, что, во-первых, будут составлены лингвистические корпусы художественных произведений, к которым планируется предоставить открытый доступ, что позволит другим научным коллективам решать с их помощью свои собственные задачи. Во-вторых, полученные данные планируется использовать в качестве иллюстративного аутентичного языкового материала для теоретических и практических дисциплин вузов в рамках лингвистических направлений подготовки.

Методами исследования являются анализ (включая автоматический и автоматизированный анализ) и моделирование (в том числе построение моделей лингвистических корпусов и моделей художественной реальности произведений).

Теоретической базой исследования послужили труды, посвященные описанию корпусного подхода (Зубов, 2006; Баркович, 2016; Комалова, 2019; Gorozhanov, Guseynova, 2020a), а также работы, раскрывающие суть интерпретации художественного произведения, включая анализ идиостиля автора (Бахтин, 1975; Ноздрина, 1997; Лотман, 2000; Болотнова, Болотнов, 2012; Баранов, Добровольский, Фатеева, 2021).

Основная часть

1. Параметры интерпретации художественных произведений

Поскольку интерпретация художественного произведения может иметь различный характер – от субъективной трактовки причин поведения персонажей до сбора исключительно статистических параметров, необходимо уточнить, к чему именно мы стремимся в рамках нашей работы.

Итак, с помощью разрабатываемого нами метода мы планируем достичь результатов, которые бы позволили, во-первых, описать параметры художественной реальности произведения, во-вторых, составить представление о персональных характеристиках героев, в-третьих, установить черты идиостиля автора.

Описание художественной реальности в рамках нашей задачи сводится прежде всего к исследованию таких категорий, как пространство, время и качество, которые могут быть представлены в различных комбинациях, например, пространство и время («хронотоп» (Бахтин, 1975; Ноздрина, 1997; Кроо, 2020)), пространство и качество (Горожанов, Гусейнова, 2021).

Составление языковых портретов персонажей требует тщательной работы над текстом художественного произведения с высокой долей «ручного» труда, касающейся маркирования реплик и внутренней речи персонажей, причем работа осложняется еще и тем, что отделить языковой портрет героя от языкового портрета автора чрезвычайно трудно. Ю. М. Лотман (2000) отмечал этот факт, говоря, в частности, о представителях европейского авангарда: «Опыт европейского авангарда убедительно свидетельствует, что чем индивидуальнее художественный язык, тем более места занимает авторская рефлексия, направленная на язык и включенная в его же структуру» (с. 161). Планируется охарактеризовать языковые портреты героев, составив предварительно подкорпусы их речи как подмножества корпуса всего художественного произведения (Potapova, Komalova, 2018, с. 84).

Что касается исследования идиостиля автора, то в современной предметно-специальной литературе мы встречаем как работы, направленные на установление идиостиля в целом (Баранов, Добровольский, Фатеева, 2021), так и попытки описать его отдельные компоненты, в рамках чего могут быть проанализированы целые текстообразующие категории или отдельные языковые явления (Тарасевич, 2014; Соколова, Степанова, 2019; Бойчук, Джонсон, 2020; Горожанов, 2021). Мы планируем применить здесь принцип «от простого к сложному», двигаясь от отдельных языковых явлений к текстообразующим категориям.

2. Тип лингвистического корпуса

По нашему мнению, для интерпретации произведений художественной литературы целесообразно использовать сбалансированные корпусы, которые будут включать, например, тексты всех или определенных работ того или иного писателя, а возможно, даже тексты произведений писателей одного литературного направления. Такого рода корпусы применяются для решения специализированных задач и не имеют универсальных жестких требований к объему. Практика показывает, что они могут насчитывать от нескольких сотен до нескольких сотен миллионов словоупотреблений (Комалова, 2019, с. 115-116).

В итоге мы останавливаемся на создании собственного сбалансированного (специального) корпуса объемом не менее одного полного текста художественного произведения (с опцией выделения подкорпусов речи персонажей).

Итак, для первого этапа наших исследований мы выбираем *письменный одноязычный* (русский, немецкий, английский языки) *литературный художественный исследовательский статический неразмеченный полнотекстовый синхронический* корпус. Этот тип корпуса мы будем считать основным, поскольку для решения отдельных частных задач необходимо будет прибегать к помощи национальных корпусов затрагиваемых языков.

Прокомментируем некоторые из обозначенных признаков основного типа корпуса. Выбор языка корпуса зависит от языка оригинала художественного произведения. Привлечение параллельных корпусов не исключается, например, с целью сопоставительного анализа оригинального текста и его перевода (Крапивин, Степанова, 2021, с. 62). Признак «исследовательский» ставится в оппозицию к признаку «иллюстративный» (Зубов, 2006, с. 24). Под полнотекстовостью корпуса подразумевается включение в него всего текста исследуемых художественных произведений. Далее, корпус является статическим и синхроническим, поскольку заключает в себе тексты конкретных произведений художественной литературы, которые, очевидно, уже не подлежат изменениям.

Самым трудным параметром является наличие или отсутствие разметки корпуса. С одной стороны, размеченный корпус предоставляет точные сведения о явлениях, включенных в его разметку. С другой стороны, неразмеченный корпус в совокупности с эффективными приложениями обработки естественного языка (англ. natural language processing) позволяет оперативно апробировать гипотезу исследования на материале различных языков и художественных направлений.

По совокупности аргументов мы склоняемся к выбору неразмеченного корпуса, по крайней мере на начальном этапе нашего исследования, что не исключает дальнейшего перехода к размеченному корпусу. Предполагается, что результаты работы с неаннотированными текстовыми массивами могут послужить основой для определения модели аннотированного корпуса на последующих этапах исследования.

3. Программные решения

Проблеме выбора программных решений, то есть по сути чисто технологическому вопросу, мы придаем особое значение, т.к. от этого зависит не только скорость обработки данных, что немаловажно при объемах современных лингвистических корпусов, но и сама достоверность получаемых результатов.

Мы остановимся на языке программирования Python, который на текущий момент является мировым лидером по популярности благодаря своей универсальности и наличию большого набора библиотек для решения различных прикладных задач (https://tiobe.com/tiobe-index/). Среди упомянутых библиотек пристального внимания заслуживает spaCy, которая представляет собой набор разнообразных инструментов для расширенной обработки естественного языка (Добровольский, Кротова, Цветаева и др., 2021, с. 6). Особенность spaCy состоит в том, что этот инструмент имеет встроенные базы данных 19-ти языков (включая необходимые нам русский, немецкий и английский), которые позволяют получать не только статистические данные текста, но также составлять полноценные конкордансы по заданным шаблонам (Ayre, Bittar, Kam et al., 2021; Okhapkin, Okhapkina, Iskhakova et al., 2021). Другими возможностями библиотеки являются выделение в тексте имен собственных и различных частей речи и членов предложения, а также автоматическое резюмирование (Jugran, Kumar, Tyagi et al., 2021). Этот факт, а также то, что spaCy можно самостоятельно «обучать», придают ей некоторые черты надкорпусной базы данных (Гончаров, Инькова, 2021). Не удивительно, что spaCy характеризуют как самую популярную библиотеку для обработки текстовых массивов (Kozhevnikov, Pankratova, 2020, с. 316).

Таким образом, сегодня в распоряжении корпусного лингвиста имеются мощные инструменты, позволяющие достаточно эффективно работать с текстами без их предварительной разметки. Предполагается встраивать модули spaCy в комплекты программ, разработанные нами в рамках развития концепции профессионально ориентированного программирования (Gorozhanov, Guseynova, 2020b).

4. Предварительная апробация метода

Для предварительной апробации метода был выбран роман Ф. Кафки «Замок», оригинальный текст которого составил сбалансированный неразмеченный лингвистический корпус.

Корпус представляет собой файл TXT, программно разбитый на предложения, каждое из которых начинается с новой строки. Название романа, названия глав, а также нумерация страниц удалены.

Токенизатор spaCy выделил в корпусе 136410 элементов (токенов), к которым были отнесены не только словоформы, но также знаки препинания и знаки новой строки. В автоматическом режиме мы выделили так называемые «сущности», т.е. имена собственные и локации в порядке упоминания от самого частого к самому редкому (своего рода ономастическая модель или художественный ономастикон романа (Косиченко, 2017, с. 14)).

Результат работы программы нетривиален, поскольку spaCy проводит не простую частотную выборку и сравнение токенов с имеющимися в ее базе данных образцами, но и анализирует их окружение. После ручного объединения различных падежных форм одной и той же леммы был получен следующий список персоналий (более 10-ти раз упоминаются в полученном результате): K. – 724 раза, Frieda – 296 раз, Barnabas – 177 раз, Klamm – 120 раз, Amalia – 88 раз, Hans Brunswick – 83 раза, Olga – 70 раз, Pepi – 53 раза, Bürgel – 33 раза, Jeremias – 30 раз, Sordini – 24 раза, Sortini – 21 раз, Gerstäcker – 15 раз, Erlanger – 15 раз, Lasemann – 11 раз.

Зная содержание романа, можно сказать, что в пропорции количественных показателей достаточно точно отражается важность тех или иных героев. С другой стороны, программа не смогла выделить существительное “Gehilfen” (подмастерья), т.е. Jeremias и Artur вместе. При возвращении к идее рассмотрения языкового портрета персонажей очевидным становится тот факт, что частотные герои романа обязательны для такого рода анализа.

Среди значимых локаций романа программа зафиксировала следующие (в алфавитном порядке): Brükkenhof (постоялый двор «У моста»), Dorf (деревня), Erde (Земля), Fenster (окно), Haus (дом), Herrenhof (господский постоялый двор), Kirche (церковь), Schloss (замок). Мы не приводим здесь частотные показатели по причине значительного отличия полученного значения и фактического количества данных словоформ в тексте (например, “Herrenhof” встречается в романе 57 раз, когда как spaCy обозначила это существительное как локацию всего 12 раз). Тем не менее, упомянутые локации действительно являются ключевыми для организации художественной реальности романа. Программа «упустила» всего несколько важных локаций, например, Zimmer (комната), Ausschank (~барная стойка), Schule (школа) (Горожанов, Гусейнова, 2021, с. 27).

Интересны данные о временной структуре романа, отраженной в глагольных формах. Расчет показывает, что глаголы, имеющие в spaCy метку «настоящее», употребляются в романе 6928 раз, а глаголы с меткой «прошедшее» – 8028 раз. Интерпретировать этот результат можно различным образом. Например, такое большое количество форм настоящего времени объясняется тем, что в романе значительное место занимают диалоги персонажей, тогда как в прошедшем времени приводятся в основном различного рода описания. Для нас важно то, что этот результат был получен быстро и без предварительной разметки корпуса художественного произведения.

Заключение

Итак, мы сформулировали положения метода интерпретации художественного произведения в русле корпусного подхода и провели его первичную апробацию на конкретном языковом материале. Мы пришли к выводу о том, что главными принципами метода явились: использование неразмеченного лингвистического корпуса, который впоследствии сможет стать базой для аннотированной версии; применение в качестве программного инструмента языка Python с привлечением библиотеки spaCy. Предлагаемый метод можно считать универсальным для избранных языков: русского, немецкого и английского, поскольку различия между этими языками в нашем случае не являются критическими.

Перспективой дальнейшего исследования является уточнение положений предлагаемого метода, а также составление трех сбалансированных лингвистических корпусов к романам Ф. Кафки «Замок», «Америка» и «Процесс», одного лингвистического корпуса по сериям рассказов Дж. Лондона «Смок Беллью» и «Смок и малыш» и лингвистических корпусов текстов перевода указанных произведений на русский язык с целью проведения сопоставительных исследований.

Источники | References

1. Баранов А. Н., Добровольский Д. О., Фатеева Н. А. Идиостиль Ф. М. Достоевского: направления изучения // Вестник Российского университета дружбы народов. Серия «Теория языка. Семиотика. Семантика». 2021. Т. 12. № 2. DOI: 10.22363/2313-2299-2021-12-2-374-389
2. Баркович А. А. Корпусная лингвистика: специфика современных метаописаний языка // Вестник Томского государственного университета. 2016. № 406. DOI: 10.17223/15617793/406/1
3. Бахтин М. М. Формы времени и хронотопа в романе. Очерки по исторической поэтике // Бахтин М. М. Вопросы литературы и эстетики. М.: Худож. лит., 1975.
4. Бойчук Е. И., Джонсон М. А. Испанские количественные наречия в системе элементов авторского идиолекта (на материале романов XIX-XX вв.) // Филологические науки. Вопросы теории и практики. 2020. Т. 13. Вып. 5. DOI: 10.30853/filnauki.2020.5.48
5. Болотнова Н. С., Болотнов А. В. Когнитивный стиль языковой личности в структуре модели идиостиля: к постановке проблемы // Сибирский филологический журнал. 2012. № 4.

6. Гончаров А. А., Инькова О. Ю. Извлечение знаний о средствах выражения логико-семантических отношений при помощи надкорпусной базы данных // Информатика и ее применения. 2021. Т. 15. № 2. DOI: 10.14357/19922264210214

7. Горожанов А. И. Особенности употребления модальных глаголов в романе Ф. Кафки «Замок» // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2021. № 3 (845). DOI: 10.52070/2542-2197_2021_3_845_44

8. Горожанов А. И., Гусейнова И. А. Прикладные аспекты анализа и интерпретации текстов (на материале немецкого и русского языков). Казань: Бук, 2021.

9. Добровольский Д. О., Кротова Е. Б., Цветаева Е. Н., Шарандин А. В. Служебные и дискурсивные слова: вариативность грамматической нормы // Германистика 2021: nove et nova: мат. IV Междунар. науч. конф. (г. Москва, 10-12 ноября 2021 г.). М.: МГЛУ, 2021.

10. Зубов А. В. Корпусная лингвистика: возможности и перспективы // Русский язык: система и функционирование (к 80-летию профессора П. П. Шубы): мат. III Междунар. науч. конф. (г. Минск, 6-7 апреля 2006 г.): в 2-х ч. Мн.: РИВШ, 2006. Ч. 1.

11. Козлова Н. В. Лингвистические корпуса: определение основных понятий и типология // Вестник Новосибирского государственного университета. Серия «Лингвистика и межкультурная коммуникация». 2013. Т. 11. № 1.

12. Комалова Л. Р. Корпусные исследования в лингвистике: письменный текст // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Серия 6 «Языкознание». Реферативный журнал. 2019. № 3.

13. Косиченко Е. Ф. Имя собственное в семиотическом пространстве культуры и художественного текста. М.: МГЛУ, 2017.

14. Крапивин Ю. Б., Степанова Д. В. Решение задачи автоматического перевода научно-технических текстов с английского языка на белорусский // Вестник Брестского государственного технического университета. 2021. № 1 (124). DOI: 10.36773/1818-1212-2021-124-1-61-65

15. Кроо К. Хронотопная динамика в стихотворении М. Ю. Лермонтова «Выхожу один я на дорогу...» // Известия Уральского федерального университета. Серия 2 «Гуманитарные Науки». 2020. Т. 22. № 1 (196). DOI: 10.15826/izv2.2020.22.1.012

16. Лотман Ю. М. Семиосфера. СПб.: Искусство-СПБ, 2000.

17. Ноздрина Л. А. Взаимодействие грамматических категорий в художественном тексте: на материале немецкого языка: дисс. ... д. филол. н. М., 1997.

18. Соколова Е. Н., Степанова Ю. В. Языковая реализация авторской модальности в романе М. А. Булгакова «Мастер и Маргарита» // Вопросы когнитивной лингвистики. 2019. № 2. DOI: 10.20916/1812-3228-2019-2-103-112

19. Солнышкина М. И., Гатиятуллина Г. М. История развития корпусной лингвистики (на примере англоязычных корпусов) // Вестник Томского государственного университета. Филология. 2020. № 63. DOI: 10.17223/19986645/63/8

20. Тарасевич Л. А. Семантика и функционирование пространственных предлогов (на материале немецкого и русского языков). Мн.: Минский гос. лингв. ун-т, 2014.

21. Яновская И. В., Чижикова О. В., Золотых Н. В. Лингвокогнитивные механизмы индивидуально-авторского начала в поэтическом дискурсе // Вестник Волгоградского государственного университета. Серия 2 «Языкознание». 2019. Т. 18. № 3. DOI: 10.15688/jvolsu2.2019.3.21

22. Ayre K., Bittar A., Kam J., Verma S., Howard L. M., Dutta R. Developing a Natural Language Processing Tool to Identify Perinatal Self-Harm in Electronic Healthcare Records // PLoS ONE. 2021. Vol. 16. Iss. 8. DOI: 10.1371/journal.pone.0253809

23. Bolshina A. S., Loukachevitch N. V. Generating Training Data for Word Sense Disambiguation in Russian // Komp'juternaja Lingvistika i Intellektual'nye Tehnologii. 2020. № 19. DOI: 10.28995/2075-7182-2020-19-119-132

24. Gorozhanov A. I., Guseynova I. A. Korpusanalyse der Konstituenten Grammatischer Kategorien im Literarischen Text mit Berücksichtigungder Linguoregionalen Komponente // Журнал Сибирского федерального университета. Серия «Гуманитарные науки». 2020a. № 13 (12). DOI: 10.17516/1997-1370-0702

25. Gorozhanov A. I., Guseynova I. A. Programming for Specific Purposes in Linguistics: A New Challenge for the Humanitarian Curricula // Training, Language and Culture. 2020b. Vol. 4. Iss. 4. DOI: 10.22363/2521-442X-2020-4-4-23-38

26. Jugran S., Kumar A., Tyagi B. S., Anand V. Extractive Automatic Text Summarization Using SpaCy in Python NLP // 2021 International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2021. Greater Noida, 2021. DOI: 10.1109/ICACITE51222.2021.9404712

27. Kim M., Kwon H. Word Sense Disambiguation Using Prior Probability Estimation Based on the Korean Wordnet // Electronics. 2021. Vol. 10. Iss. 23. DOI: 10.3390/electronics10232938

28. Kozhevnikov V. A., Pankratova E. S. Research of Text Pre-Processing Methods for Preparing Data in Russian for Machine Learning // Theoretical & Applied Science. 2020. Vol. 4. Iss. 84. DOI: 10.15863/TAS.2020.04.84.55

29. Okhapkin V. P., Okhapkina E. P., Iskhakova A. O., Iskhakov A. Y. Constructing of Semantically Dependent Patterns Based on SpaCy and StanfordNLP Libraries // Communications in Computer and Information Science. 2021. Vol. 1395. DOI: 10.1007/978-981-16-1480-4_45

30. Potapova R. K., Komalova L. R.Russian SND Full-Text Annotated Polylogues Database: "Political Transformations" Sub-Corpus // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2018. № 6 (797).

Информация об авторах | Author information

RU

Горожанов Алексей Иванович[1], д. филол. н., доц.
Степанова Дарья Валерьевна[2], к. филол. н., доц.
[1] Московский государственный лингвистический университет
[2] Минский государственный лингвистический университет, Республика Беларусь

EN

Gorozhanov Alexey Ivanovich[1], Dr
Stepanova Darya Valeryevna[2], PhD
[1] Moscow State Linguistic University
[2] Minsk State Linguistic University, The Republic of Belarus

[1] *a_gorozhanov@mail.ru,* [2] *daryastepanova79@gmail.com*

Информация о статье | About this article

Дата поступления рукописи (received): 13.12.2021; опубликовано (published): 31.01.2022.

Ключевые слова (keywords): интерпретация художественного произведения; корпусный подход; идиостиль; Ф. Кафка; work of fiction interpretation; corpus approach; idiostyle; spaCy; F. Kafka.