



Архитектура сбалансированного лингвистического корпуса, полученного автоматическим путем (опыт Московского государственного лингвистического университета)

А. И. Горожанов

Московский государственный лингвистический университет, Москва, Россия
a.gorozhanov@linguanet.ru

Аннотация. Цель настоящего прикладного исследования – продемонстрировать возможности современных программных решений построения сбалансированного лингвистического корпуса на основе процедур обработки естественного языка, применяемых в лаборатории фундаментальных и прикладных проблем виртуального образования Московского государственного лингвистического университета. В ходе работы применяются описательный метод, а также методы моделирования и прогнозирования. В качестве материала исследования выступает авторский программный комплекс «Генератор сбалансированного лингвистического корпуса и корпусный менеджер». В результате описаны новые функции программного комплекса и обозначена перспектива его развития в виде двух параллельных направлений.

Ключевые слова: прикладная лингвистика, корпусная лингвистика, корпусный менеджер, сбалансированный лингвистический корпус, архитектура программного обеспечения, реляционная база данных, отечественное программное обеспечение, обработка естественного языка

Для цитирования: Горожанов А. И. Архитектура сбалансированного лингвистического корпуса, полученного автоматическим путем (опыт Московского государственного лингвистического университета) // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2024. Вып. 11 (892). С. 24–30.

Original article

Architecture of a Balanced Linguistic Corpus Built Automatically (experience of Moscow State Linguistic University)

Alexey I. Gorozhanov

Moscow State Linguistic University, Moscow, Russia
a.gorozhanov@linguanet.ru

Abstract. The purpose of this applied research is to demonstrate the capabilities of modern software solutions for constructing a balanced linguistic corpus based on natural language processing procedures used at the Laboratory for Fundamental and Applied Issues of Virtual Education at Moscow State Linguistic University. The descriptive method, as well as modeling and forecasting methods, are used during the study. The material of the research is the author's software package "Balanced Linguistic Corpus

Generator and Corpus Manager". As a result, new functions of the software package are described and the prospects for its development in the form of two parallel directions are outlined.

Keywords: applied linguistics, corpus linguistics, corpus manager, balanced linguistic corpus, software architecture, relational data base, domestic software, natural language processing

For citation: Gorozhanov, A. I. (2024). Architecture of a balanced linguistic corpus built automatically (experience of Moscow State Linguistic University). *Vestnik of Moscow State Linguistic University. Humanities*, 11(892), 24–30. (In Russ.)

ВВЕДЕНИЕ

В 2023 году в лаборатории фундаментальных и прикладных проблем виртуального образования Московского государственного лингвистического университета началась экспериментальная разработка программных продуктов для автоматической генерации и оперирования сбалансированными лингвистическими корпусами, которая завершилась регистрацией первой устойчивой версии программного комплекса¹.

Лаборатория показала себя как подразделение «полного цикла», т. е. она в полной мере участвует в планировании, разработке, апробации, регистрировании и коммерциализации программного продукта.

На сегодняшний день можно констатировать, что комплекс расширился дополнительными функциями, сохраняя при этом возможность полностью автоматической разметки исходного текста. Отсюда вытекает гипотеза нашего исследования, которая заключается в том, что на момент планирования разработчиком была выбрана удачная архитектура лингвистического корпуса, обладающая параметрами гибкости (возможность модификации в течение жизненного цикла) и расширяемости (добавления новых функций без изменения основных параметров).

Заметим, что под *архитектурой* мы понимаем здесь именно *архитектуру программного обеспечения*, т. е. набор ключевых решений об организации программной системы.

Задачами исследования являются следующие:

- указать основные параметры базы данных сбалансированного лингвистического корпуса в ее актуальном состоянии (описательный метод и метод моделирования);

- перечислить возможности текущей версии корпусного менеджера, оперирующего этой базой данных (описательный метод);
- обозначить перспективу развития программного комплекса, особенно варианты привлечения ручной и автоматизированной разметки (метод прогнозирования).

Актуальность исследования подтверждается стабильно высоким интересом к проблемам прикладной и корпусной лингвистики в научных публикациях, посвященных анализу публицистического дискурса [Бондарчук, 2024; Красикова, 2024; Степанова, 2023], юмористического дискурса [Гусейнова, Косиченко, 2024; Соколова, Голубкова, 2024], образовательного дискурса [Котюрова, Щеголева, 2024; Куприянов, Солнышкина, Лехницкая, 2023].

Новизна работы обусловлена тем, что впервые дается систематическое описание указанного программного комплекса в текущей версии.

С точки зрения фундаментальной науки работа вносит вклад в теорию проектирования программного обеспечения лингвистического назначения на основе элементов технологий искусственного интеллекта (обработки естественного языка).

Практическая ценность исследования заключается в том, что предложенное осмысление и прогнозирование конкретных программных решений позволит значительно улучшить текущую версию программного комплекса, что, в свою очередь, повысит достоверность результатов поисковых запросов и качество реализованных с его помощью экспериментов.

ПАРАМЕТРЫ БАЗЫ ДАННЫХ СБАЛАНСИРОВАННОГО ЛИНГВИСТИЧЕСКОГО КОРПУСА

База данных, используемая в этом и многих других наших исследованиях, является реляционной и состоит из двух таблиц: таблицы токенов и таблицы предложений. Это позволяет получать в ходе поисковых запросов контексты на уровне предложений, в которых встречается то или иное языковое

¹Свидетельство о государственной регистрации программы для ЭВМ № 2023683209 Российская Федерация. «Генератор сбалансированного лингвистического корпуса и корпусный менеджер»: № 2023682269; заявл. 25.10.2023; опублик. 03.11.2023 / А. И. Горозханов; заявитель федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный лингвистический университет». EDN JHFUVU.

Таблица 1

ВИЗУАЛИЗАЦИЯ ТАБЛИЦЫ ПРЕДЛОЖЕНИЙ С ЗАПОЛНЕННЫМИ ДОПОЛНИТЕЛЬНЫМИ ЯЧЕЙКАМИ

	id	sentnum	senttext	sentoption01	sentoption02	sentoption03	sentoption04	sentoption05
	F...	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	1	1	It was a bright...	3	1	2	-1.0	3.0
2	2	2	Winston Smith...	4	1	3	-2.0	4.0
3	3	3	The hallway s...	2	0	2	-2.0	2.0
4	4	4	At one end of i...	0	0	0	0.0	0.0
5	5	5	It depicted si...	1	1	0	1.0	1.0
6	6	6	Winston made...	0	0	0	0.0	0.0
7	7	7	It was no use ...	0	0	0	0.0	0.0
8	8	8	Even at the be...	2	2	0	2.0	2.0
9	9	9	It was part of ...	0	0	0	0.0	0.0
10	10	10	The flat was s...	2	1	1	0.0	2.0
11	11	11	On each landin...	0	0	0	0.0	0.0

Таблица 2

ВИЗУАЛИЗАЦИЯ ТАБЛИЦЫ ТОКЕНОВ
С ДВУМЯ ЗАПОЛНЕННЫМИ ДОПОЛНИТЕЛЬНЫМИ ЯЧЕЙКАМИ ИЗ ПЯТИ

	id	tokennum	sent_num	tokentext	tokenpos	tokenlemma	tokenattr	tokenoption01	tokenoption02	tokenoption03	tokenoption04	tokenoption05
	Fil...	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter	Filter
1	1	0	1	It	PRON	it	Case=Nom ...	NULL	1	NULL	NULL	NULL
2	2	1	1	was	AUX	be	Mood=Ind ...	NULL	1	NULL	NULL	NULL
3	3	2	1	a	DET	a	Definite=In...	NULL	5	NULL	NULL	NULL
4	4	3	1	bright	ADJ	bright	Degree=Pos	1	5	NULL	NULL	NULL
5	5	4	1	cold	ADJ	cold	Degree=Pos	-1	5	NULL	NULL	NULL
6	6	5	1	day	NOUN	day	Number=Sing	NULL	1	NULL	NULL	NULL
7	7	6	1	in	ADP	in		NULL	5	NULL	NULL	NULL
8	8	7	1	April	PROPN	April	Number=Sing	NULL	6	NULL	NULL	NULL
9	9	8	1	,	PUNCT	,	PunctType=...	NULL	1	NULL	NULL	NULL
10	10	9	1	and	CCONJ	and	ConjType=C...	NULL	1	NULL	NULL	NULL
11	11	10	1	the	DET	the	Definite=De...	NULL	11	NULL	NULL	NULL
12	12	11	1	clocks	NOUN	clock	Number=Plur	NULL	13	NULL	NULL	NULL
13	13	12	1	were	AUX	be	Mood=Ind T...	NULL	13	NULL	NULL	NULL
14	14	13	1	striking	VERB	strike	Aspect=Pro...	-1	1	NULL	NULL	NULL
15	15	14	1	thirteen	NUM	thirteen	NumType=C...	NULL	13	NULL	NULL	NULL
16	16	15	1	.	PUNCT	.	PunctType=...	NULL	13	NULL	NULL	NULL

явление. Здесь мы можем соотнести употребляемые нами понятия «контекст» / «предложение» с понятием «конкорданс», распространенным в корпусной лингвистике [Гик, 2024; Бобунова, 2023].

Однако выбор предложения в качестве «старшего токена»¹ позволяет, по нашему мнению, работать

¹По отношению к «младшему токenu» – словоформе, числу или знаку пунктуации.

не просто с последовательностью токенов, а со смысловой единицей [Ким, 2021].

Под сбалансированным лингвистическим корпусом мы понимаем лингвистический корпус, отличный от национального лингвистического корпуса и призванный отразить характер употребления языковых явлений, свойственных не национальному языку вообще, а языку отдельного человека, художественного произведения, новостного издания и пр. [Горожанов, Степанова, 2022].

Приведем примеры графического изображения обеих таблиц базы данных (см. табл. 1 и 2)¹. В приведенных выше вариантах в таблице токенов заполнены дополнительные колонки 1 и 2. В первой расположены данные тонального словаря, если токен имеет положительную (1) или отрицательную (-1) коннотацию [Горожанов, 2023, с. 3889–3891]. Во второй помещаются так называемые зависимости, т. е. числовое указание на номер токена, который является главным для текущего токена [Горожанов, 2024а, с. 184].

В таблице предложений заполнены все дополнительные параметры, полученные расчетами по данным первой дополнительной колонки токенов. Это коннотативная плотность (ρ_{con}) предложения (общее количество положительно и отрицательно коннотированных токенов); сумма положительных токенов; сумма отрицательных токенов; сумма значений коннотированных токенов; коннотативная амплитуда (A_{con}) предложения (сумма значений всех коннотированных токенов по модулю) [Горожанов, 2023, с. 3890].

ВОЗМОЖНОСТИ ТЕКУЩЕЙ ВЕРСИИ КОРПУСНОГО МЕНЕДЖЕРА

В предыдущих публикациях мы описали базовые функции нашего корпусного менеджера:

- 1) загрузку лингвистического корпуса;
- 2) вывод списка частотности лемм;
- 3) вывод предложения по заданному номеру;
- 4) вывод всех предложений, содержащих заданную лемму;
- 5) вывод всех предложений, содержащих заданные леммы (по списку);
- 6) вывод всех предложений, содержащих заданный токен;
- 7) вывод всех предложений, содержащих заданные токены (по списку);
- 8) вывод всех предложений, содержащих заданную часть речи или несколько частей речи;

¹В качестве языкового материала здесь выступает содержание сбалансированного лингвистического корпуса текста цикла рассказов Дж. Лондона «Смок Белью. Смок и Малыш» в оригинальном изложении [Горожанов, Гусейнова, Степанова, 2022].

- 9) вывод всех предложений, содержащих токены с заданными морфологическими характеристиками²;
- 10) вывод всех предложений, содержащих токены с заданными значениями морфологических характеристик (например, грамматический род = «мужской», залог = «активный» и пр.);
- 11) вывод всех предложений, отвечающих запросу SQL (например, `SELECT *FROM tokens WHERE id < 100`);
- 12) вывод по сложному запросу (учет совокупности характеристик, например, заданный токен + часть речи «существительное» + число «множественное» и пр.) [Горожанов, Гусейнова, Степанова, 2024].

На сегодняшний день программный комплекс расширен как по линии генератора, так и по линии корпусного менеджера. Генератор получил важную функцию разметки зависимости токенов друг от друга (см. табл. 2).

Корпусный менеджер приобрел целый набор расширений:

- 1) объединение нескольких корпусов в один;
- 2) получение частотного списка по заданным частям речи;
- 3) поиск фразеологизмов по заданным параметрам [Горожанов, 2024б];
- 4) поиск по регулярным выражениям³;
- 5) поиск по последовательности символов в начале, середине и в конце токена;
- 6) модуль оценки контекста заданной леммы (при условии наличия разметки зависимостей токенов в базе данных);
- 7) экспериментальный модуль генерации учебного контента для LMS Moodle;
- 8) поиск по *специальному ручному запросу*, который позволяет находить последовательности токенов с заданными характеристиками (например, «артикле + прилагательное + существительное», «артикле + 2 любых токена + существительное» и т. п.)

ПЕРСПЕКТИВА РАЗВИТИЯ ПРОГРАММНОГО КОМПЛЕКСА

Заявленный программный комплекс, на наш взгляд, имеет два пути развития, которые могут осуществляться параллельно. *Во-первых*, может быть продолжена работа над расширением функций

²Под «морфологическими характеристиками» здесь, согласно терминологии области обработки естественного языка, понимаются скорее грамматические признаки (грамматический род, падеж, залог и пр.).

³URL: <https://docs.python.org/3/library/re.html>

автоматического получения новых данных, т. е. без участия оператора / исследователя. В этой связи перспективной кажется работа по генерации тестовых заданий на различные языковые явления по разным языкам. *Во-вторых*, важной является ветка развития по созданию вспомогательных банков данных, которые позволят решать специализированные задачи. В качестве таких банков данных могут выступить тональные словари для определенных типов дискурса. Потребность в подобной специализации обусловлена особенностями подязыка отдельно взятой предметной области [Писарик, 2021].

Отметим также и формат технической реализации нашего программного продукта, который сознательно выполнен как отдельное приложение, существующее в виде инсталляции. Использование базы данных типа SQL (в базовой версии SQLite) позволяет осуществить миграцию на MySQL при условии перевода программного ядра с Python, например, на PHP, что в результате позволит создать корпусный менеджер в виде веб-сервиса – возможно, и с некоторым изменением функциональных возможностей. Тем не менее формально такое решение является осуществимым ввиду изначального применения перспективных и универсальных технологий.

ЗАКЛЮЧЕНИЕ

Итак, мы описали базовые параметры программного комплекса «Генератор сбалансированного лингвистического корпуса и корпусный менеджер», приведя описание реляционной базы данных, состоящей из двух таблиц.

Гипотеза исследования, которая заключалась в том, что на момент планирования разработчиком была выбрана удачная архитектура лингвистического корпуса, обладающая параметрами гибкости (возможность модификации в течение жизненного цикла) и расширяемости (добавления новых функций без изменения основных параметров), может считаться подтвержденной.

Мы указали на возможности текущей версии корпусного менеджера, оперирующего этой базой данных, перечислив новые функции, которые были детально описаны в научных публикациях.

Далее мы обозначили перспективу развития программного комплекса, причем, в нашем понимании, оно может идти по параллельным траекториям: в русле развития автоматических универсальных процедур и в русле развития автоматизированных решений, которые предусматривают участие специально созданных банков данных и корректировку получаемых данных человеком.

СПИСОК ИСТОЧНИКОВ

1. Бондарчук Г. Г. Семиотические функции английских наименований одежды в публицистическом тексте (корпусное исследование) // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2024. Вып. 4 (885). С. 23–29. EDN BXILCR.
2. Красикова Е. А. Роль корпусного менеджера в анализе употребления имен собственных в текстах электронных СМИ (на примере англоязычного корпуса CNN) // Филологические науки в XXI веке: актуальность многополярности перспективы развития: сборник научных трудов. Краснодар: Кубанский государственный университет 2024. С. 45–49. EDN JPRHAE.
3. Степанова Д. В. Программный комплекс для генерации динамического корпуса текстов СМИ // Вестник Минского государственного лингвистического университета. Серия 1: Филология. 2023. № 6 (127). С. 123–130. EDN FMBTKO.
4. Гусейнова И. А., Косиченко Е. Ф. Грани смешного и юмор без границ: семиотика комических текстов разных жанров. Казань: Бук 2024. ISBN 978-5-907839-92-2. EDN PSLMFL.
5. Соколова В. Л., Голубкова Е. Е. Дискурсивный механизм и концептуальные основания лингвостилистической кластеризации в англоязычных коротких шутках // Когнитивные исследования языка. 2024. № 2–2 (58). С. 215–218. EDN OHNINL.
6. Котюрова И. А., Щеголева Л. В. Визуализация образовательных данных в немецкоязычном корпусе студенческих текстов // Перспективы науки и образования. 2024. № 2 (68). С. 578–594. DOI 10.32744/pse.2024.2.35. EDN UTDLFM.
7. Куприянов Р. В., Солнышкина М. И., Лехницкая П. А. Параметрическая таксономия учебных текстов // Вестник Волгоградского государственного университета. Серия 2: Языкознание. 2023. Т. 22. № 6. С. 80–94. DOI 10.15688/jvolsu2.2023.6.6. EDN VFCVLW.
8. Гик А. В. Приложения к Конкордансу М. Кузмина // Труды института русского языка им. В. В. Виноградова. 2024. № 1. С. 227–243. DOI 10.31912/pvrli-2024.1.22. EDN NVKTQL.

9. Бобунова М. А. Об исследовательском потенциале лексикографических комплексов фольклорных текстов // Вопросы лексикографии. 2023. № 28. С. 44–65. DOI 10.17223/22274200/28/3. EDN SFNPOP.
10. Ким И. Е. Пунктуация «говорящего» и пунктуация «слушающего»: ономаσιологический и семасиологический подход в пунктуации // Труды института русского языка им. В. В. Виноградова. 2021. № 3. С. 252–260. DOI 10.31912/pvrl-2021.3.20. EDN BZDVOQ.
11. Горожанов А. И., Степанова Д. В. Интерпретация художественного произведения: корпусный подход // Филологические науки. Вопросы теории и практики. 2022. Т. 15. № 1. С. 203–208. DOI 10.30853/phil20220020. EDN TCZLAF.
12. Горожанов А. И., Гусейнова И. А., Степанова Д. В. Стандартизированная процедура получения статистических параметров текста (на материале цикла рассказов Дж. Лондона «Смок Белью. Смок и Малыш») // Вестник Минского государственного лингвистического университета. Серия 1: Филология. 2022. № 4 (119). С. 7–13. EDN PXAVUX.
13. Горожанов А. И. Расширение стандартного сбалансированного лингвистического корпуса построенного по правилам sраСу коннотативными характеристиками // Филологические науки. Вопросы теории и практики. 2023. Т. 16. № 11. С. 3888–3893. DOI 10.30853/phil20230594. EDN FVUIUL.
14. Горожанов А. И. Метод программного анализа контекста лексической единицы // Актуальные проблемы филологии и педагогической лингвистики. 2024а. № 3. С. 178–190. DOI 10.29025/2079-6021-2024-3-178-190. EDN QNFATA.
15. Горожанов А. И., Гусейнова И. А., Степанова Д. В. Обработка естественного языка и художественный текст: база для корпусного исследования // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2024. Т. 15. №1. С. 195210. DOI 10.22363/2313-2299-2024-15-1-195-210.
16. Горожанов А. И. Алгоритмы поиска фразеологизмов в лингвистическом корпусе с морфологической разметкой (индоевропейские языки) // Филологические науки. Вопросы теории и практики. 2024б. Т. 17. № 1. С. 132–138. DOI 10.30853/phil20240020. EDN JTWSIQ.
17. Писарик О. И. Принципы разработки базы данных подъязыка предметной области «строительство» // Вестник Московского государственного лингвистического университета. Гуманитарные науки. 2021. № 5 (847). С. 150–160. DOI 10.52070/2542-2197_2021_5_847_150. EDN RKPNSU.

REFERENCES

1. Bondarchuk, G. G. (2024). Semiotic functions of English clothing names in a journalistic text (corpus-based study). Vestnik of Moscow State Linguistic University. Humanities, 4(885), 23–29. EDN BXILCR. (In Russ.)
2. Krasikova, E. A. (2024). The role of the corpus manager in analyzing the use of proper names in electronic media texts (on the material of the English-speaking CNN corpus). Filologicheskie nauki v XXI veke: aktual'nost' mnogopolyarnost' perspektivy razvitiya (pp. 45–49): collection of scientific papers. Krasnodar: Kuban State University. EDN JPRHAE. (In Russ.)
3. Stepanova, D. V. (2023). Software package for generating a dynamic media texts corpus. Minsk State Linguistic University Bulletin. Series 1. Philology, 6(127), 123–130. EDN FMBTKO. (In Russ.)
4. Guseynova, I. A., Kosichenko, E. F. (2024). Grani smeshnogo i yumor bez granits: semiotika komicheskikh tekstov raznykh zhanrov = The Facets of the Funny and Humor Without Borders: Semiotics of Comic Texts of Different Genres. Kazan: Buk. ISBN 978-5-907839-92-2. EDN PSLMFL. (In Russ.)
5. Sokolova, V. L., Golubkova, E. E. (2024). Discursive mechanism and conceptual foundations of shaping linguostylistic clusters in the English-language one-liner jokes. Cognitive studies of language, 2–2(58), 215–218. EDN OHNINL. (In Russ.)
6. Kotiurova, I. A., Shchegoleva, L. V. (2024). Visualization of educational data in a German-language corpus of student texts. Perspectives of science and education, 2(68), 578–594. 10.32744/pse.2024.2.35. EDN UTDLFM. (In Russ.)
7. Kupriyanov, R. V., Solnyshkina, M. I., Lekhnitskaya, P. A. (2023). Parametric Taxonomy of Educational Texts. Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie [Science Journal of Volgograd State University. Linguistics], 22(6), 80–94. 10.15688/jvolsu.2023.6.6. EDN VFCVLW. (In Russ.)
8. Gik, A. V. (2024). The appendicies to the concordance of M. Kuzmin. Proceedings of the V. V. Vinogradov Russian Language Institute, 1, 227–243. 10.31912/pvrl-2024.1.22. EDN NVKTQL. (In Russ.)
9. Bobunova, M. A. (2023). On the research potential of lexicographic complexes of folklore texts. Russian journal of lexicography, 28, 44–65. 10.17223/22274200/28/3. EDN SFNPOP. (In Russ.)

10. Kim, I. E. (2021). Punctuation of the "speaker" and punctuation of the "listener": onomasiological and the semasiological approach in punctuation. Proceedings of the V. V. Vinogradov Russian Language Institute, 3, 252–260. 10.31912/pvrli-2021.3.20. EDN BZDVOQ. (In Russ.)
11. Gorozhanov, A. I., Stepanova, D. V. (2022). Work of fiction interpretation: corpus approach. Philology. Theory & practice, 15(1), 203–208. 10.30853/phil20220020. EDN TCZLAF. (In Russ.)
12. Gorozhanov, A. I., Guseynova, I. A., Stepanova, D. V. (2022). Standardized procedure for obtaining statistical parameters of a text (on the material of the stories by J. London "Smoke Bellew. Smoke and Shorty"). Minsk State Linguistic University bulletin. Series 1. Philology, 4(119), 7–13. EDN PXAVUX. (In Russ.)
13. Gorozhanov, A. I. (2023). Extension of a standard balanced linguistic corpus built according to spacy rules by connotative characteristics. Philology. Theory & practice, 16(11), 3888–3893. 10.30853/phil20230594. EDN FVUIUL. (In Russ.)
14. Gorozhanov, A. I., Guseynova, I. A., Stepanova, D. V. (2024). Natural Language Processing and Fiction Text: Basis for Corpus Research. RUDN Journal Of Language Studies, Semiotics and Semantics, 15(1), 195–210. 10.22363/2313-2299-2024-15-1-195-210.
15. Gorozhanov, A. I. (2024a). Programming analysis of the lexical unit context. Current Issues in Philology and Pedagogical Linguistics, 3, 178–190. 10.29025/2079-6021-2024-3-178-190. (In Russ.)
16. Gorozhanov, A. I. (2024b). Algorithms for searching phraseological units in a linguistic corpus with morphological markup (Indo-European languages). Philology. Theory & practice, 17(1), 132–138. 10.30853/phil20240020. EDN JTWSIQ. (In Russ.)
17. Pisarik, O. I. (2021). Database design principles for the "construction" English sublanguage. Vestnik of Moscow State Linguistic University. Humanities, 5(847), 150–160. 10.52070/2542-2197_2021_5_847_150. EDN RKPNSU. (In Russ.)

ИНФОРМАЦИЯ ОБ АВТОРЕ

Горожанов Алексей Иванович

доктор филологических наук, доцент
 профессор кафедры грамматики и истории немецкого языка
 факультета немецкого языка
 Московского государственного лингвистического университета

INFORMATION ABOUT THE AUTHOR

Gorozhanov Alexey Ivanovich

Doctor of Philology, Associate Professor,
 Professor in the Department of German Language Grammar and History
 Faculty for German Language
 Moscow State Linguistic University

Статья поступила в редакцию	02.07.2024	The article was submitted approved after reviewing accepted for publication
одобрена после рецензирования	31.07.2024	
принята к публикации	06.08.2024	