

Научная статья

УДК 811.11-112; 001.4

doi: 10.17223/19996195/72/2

Потенциал корпусных менеджеров для анализа немецких лингводидактических терминов

Анжела Игоревна Едличко¹, Анна Викторовна Мишнова²

^{1, 2} Московский государственный университет им. М.В. Ломоносова, Москва, Россия

¹ ang299@yandex.ru

² mishnovaa@mail.ru

Аннотация. Рассматриваются возможности, связанные с применением корпусных менеджеров Sketch Engine и AntConc для отбора и анализа немецких терминов стандартизированного лингводидактического тестирования. Данная сфера активно развивается в связи с эволюцией лингводидактического знания, отражающегося в терминах как способах его кодирования. Отбор терминов из текстов является одним из основных этапов терминоведческого исследования и осложняется в рассматриваемой области консубстанциональным характером многих обозначений, т.е. их формальным совпадением с общеупотребительными словами. Консубстанциональные термины проявляют терминологические значения только в определенном контексте, что ведет к необходимости анализа их коллокаций в процессе терминологического отбора и ставит вопрос об эффективных инструментах его осуществления. Исследование проводится в русле корпусной лингвистики, предлагающей использование как общенаучных (описание, индукция, демонстрация, классификация, моделирование, сопоставление), так и частнонаучных методов (корпусного метода и метода контекстуального анализа).

Отсутствие представленности многих современных лингводидактических терминов в масштабных корпусах немецкого языка, таких как DeReKo, ведет к необходимости составления собственных корпусов, для чего используются корпусные менеджеры. Корпусные менеджеры являются инструментами компьютерной лингвистики, позволяющими оптимизировать терминоведческое исследование благодаря широкому набору функций. Функции корпусных менеджеров, рассматриваемых в исследовании (Sketch Engine, AntConc), могут включать, в зависимости от выбранного инструмента, автоматическое извлечение одно- и многословных терминов, составление тезауруса, вычисление меры дисперсии термина, выявление N-грамм, создание списка слов по заданным критериям, разработку конкорданса и списка типичных коллокаций, сопоставление коллокаций двух терминов. Для раскрытия потенциала применения данных корпусных менеджеров при отборе немецких терминов стандартизированного лингводидактического тестирования собран практический материал, включающий 16 руководств для экзаменаторов Гёте-сертификата с 2001 по 2023 г., которые представляются в качестве корпуса оцифрованных текстов объемом 379 348 токенов.

Сервисы Sketch Engine и AntConc показывают высокие результаты компьютерного анализа немецких терминов стандартизированного лингводидактического тестирования, связанные со структурированным представлением информации в виде конкорданса, автоматическим выявлением сочетаемости терминов, машинной классификацией терминов по заданным критериям (морфемам, частям речи). Инструменты могут оптимизировать процесс отбора терминов при

работе с большими объемами информации, предоставляя данные по заданным критериям. Несмотря на существенные достоинства корпусных менеджеров, они допускают ряд ошибок и неточностей, заключающихся в неправильной или неполной обработке информации: автоматическом извлечении общеупотребительных слов вместо терминов, неполном определении формы слова, неверном установлении коллокаций и семантических параметров. Это говорит о необходимости комбинирования компьютерных инструментов с методами ручного отбора терминов.

Ключевые слова: термин, терминоведение, немецкий язык, стандартизованное лингводидактическое тестирование, отбор терминов, коллокация, компьютерные инструменты, корпус, корпусный менеджер, Sketch Engine, AntConc

Для цитирования: Едличко А.И., Мишнова А.В. Применение корпусных менеджеров в процессе отбора терминов немецкого лингводидактического стандартизированного тестирования // Язык и культура. 2025. № 72. С. 33–50. doi: 10.17223/19996195/72/2

Original article

doi: 10.17223/19996195/72/2

Potencial of corpus managers in the analysis of German language teaching terms

Anzhela I. Edlichko¹, Anna V. Mishnova²

^{1,2}*Lomonosov Moscow State University, Moscow, Russia*

¹*ang299@yandex.ru*

²*mishnovaa@mail.ru*

Abstract. This article explores the potential of Sketch Engine and AntConc in extracting and analyzing German terms in the standardized language testing. This field is rapidly evolving in connection with the development of language teaching knowledge, which is reflected in terms as ways of its encoding. Term extraction is one of the main stages in terminological research and, in the area under consideration, is complicated by the consubstantial character of many designations – i.e., their formal coincidence with commonly used words. Consobstantial terms acquire specialized meanings only in specific contexts, which necessitates the analysis of their collocations during term selection and raises the issue of effective tools for this process. The research is carried out within the framework of corpus linguistics, which proposes the use of both general scientific methods (description, induction, deduction, classification, modeling, comparison) and field-specific methods (the corpus method and contextual analysis).

Many modern language teaching terms are absent from large German corpora like DeReKo. Consequently, there is a need to create custom corpora, for which corpus managers are used. Corpus managers, tools in computational linguistics, optimize terminological research through their diverse functions. The functions of the corpus managers considered in this study (Sketch Engine, AntConc) may include, depending on the chosen tool: automatic extraction of single-word and multi-word terms, thesaurus creation, dispersion calculation, N-gram identification, word list generation, concordance development, and collocation comparison. To assess the potential of these corpus managers in extracting German terms for standardized testing, a corpus was compiled from 16 Goethe certificate guidelines (2001–2023), totaling 379,348 tokens.

Sketch Engine and AntConc show high results in analyzing German terms for standardized testing, offering structured concordances, automatic collocation identification, and term classification according to given criteria (by morphemes, parts of speech). These tools can optimize the process of term selection when working with large data volumes by supplying data based on set criteria. Despite their significant advantages, corpus managers also exhibit errors and inaccuracies due to incorrect or incomplete data processing: automatic extraction of common words instead of terms, incomplete determination of word form, and incorrect identification of collocations and semantic parameters. This indicates the need to combine computer tools with manual methods of term extraction.

Keywords: term, terminology, the German language, standardized language testing, term extraction, collocation, computer tools, corpus, corpus manager, Sketch Engine, AntConc

For citation: Edlichko A.I., Mishnova A.V. The application of corpus managers in the retrieval of terms in the German standardized language testing. *Language and Culture*, 2025, 72, pp. 33–50. doi: 10.17223/19996195/72/2

Введение

Активное развитие терминоведения привело к необходимости появления новых форм представления терминологического знания, к которым З.И. Комарова относит «базы данных, базы знаний, машинные фонды и др.» [1. С. 27]. Эволюция информационно-коммуникационных технологий и искусственного интеллекта (ИИ) поставила перед терминоведением ряд задач, связанных с анализом больших массивов текстовой информации: машинным переводом специальных текстов, извлечением из них необходимой информации, автоматическим реферированием, аннотированием, классификацией, кластеризацией, рубрицированием, автоматизацией подготовки и редактирования текстов и др. [2. С. 9–13]. Для выполнения данных задач необходимо использование различных компьютерных инструментов – «программ, технологий организации и обработки данных», признанных, по определению Е.Г. Соколовой, «для моделирования функционирования языка в тех или иных условиях, ситуациях, проблемных сферах» [3. С. 15].

Целью статьи является анализ потенциала применения корпусных менеджеров Sketch Engine и AntConc для отбора терминов в сфере немецкого стандартизированного лингводидактического тестирования – одной из динамично развивающихся областей знания. Эволюция данной сферы обусловлена, с одной стороны, ростом популярности международных языковых стандартизованных экзаменов. С другой стороны, важную роль играет в целом прогресс в теории и методике обучения иностранным языкам, связанный с распространением ИИ, информационно-коммуникативных технологий, а также совершенствованием методов и подходов в преподавании. Эти процессы сопровождаются появлением терминов, обозначающих новые лингводидактические реалии, и ведут к

необходимости терминоведческих исследований как в теоретических (выявление особенностей лингводидактических терминов как разновидности терминов гуманитарных наук), так и в практических целях (перевод терминов на другие языки, разработка лексикографических источников).

Одним из основополагающих этапов изучения терминов является их отбор, который осложняется в сфере лингводидактического стандартизированного тестирования консубстанциональным характером терминов: некоторые из них (*Püfung, Sprache, Aufgabe* и др.) совпадают по форме с общеупотребительными словами. Терминологические значения консубстанциональных терминов проявляются только в определенном контексте, поэтому в процессе выборки терминов необходим анализ их коллокаций, под которыми мы понимаем, вслед за З.И. Комаровой, характерные сочетания слов, «чье появление рядом друг с другом основывается на регулярном характере взаимного ожидания и задается не грамматическими, а чисто семантическими факторами» [4. С. 583–584]. Необходимость анализа терминов с учетом их коллокаций ставит вопрос о применении эффективных инструментов, позволяющих работать со значительным объемом информации. В качестве таких инструментов выбраны корпусные менеджеры Sketch Engine и AntConc, являющиеся одними из наиболее распространенных в лингвистических исследованиях. Таким образом, в статье освещаются возможности для извлечения немецких лингводидактических терминов и анализа их коллокаций, которые предлагают исследователю корпусные менеджеры Sketch Engine и AntConc. В работе используются общенаучные методы описания, индукции, дедукции, классификации, моделирования, сопоставления, а также частнонаучные методы контекстного анализа и корпусный метод.

Актуальность исследования обусловлена необходимостью изучения терминов гуманитарных наук, в том числе терминов лингводидактики, для установления их особенностей. Термины представляют плодотворный исследовательский материал, поскольку кодируют лингводидактическое знание и позволяют глубже проникнуть в его специфику. Как отмечает З.И. Комарова, проблема выявления типов знания и их вербализации в естественных и искусственных языках является приоритетной в лингвистике и терминоведении третьего тысячелетия [1. С. 30]. Стремительный рост инструментов компьютерной лингвистики, особенно на базе ИИ, ведет к важности изучения их преимуществ и недостатков в терминоведческом исследовании. Попытка анализа потенциала использования корпусных менеджеров на материале немецких терминов лингводидактического стандартизированного тестирования предпринимается впервые, что подчеркивает новизну исследования. Анализ возможностей применения компьютерных инструментов для выборки

лингводидактических терминов может быть полезным для терминоведческих исследований в других тематических областях.

Методология исследования

В рамках исследования был собран эмпирический материал, включающий 16 руководств для экзаменаторов Гёте-института, опубликованных в период с 2001 по 2023 г. [5–7]. Выбор руководств в качестве практического материала объясняется тем фактом, что руководства, реагируя на изменения в стандартизированном лингводидактическом тестировании, периодически обновляются. Это делает их источником актуального знания, отражающегося в терминах. Практический материал был оцифрован и организован в виде корпуса (379 348 токенов). Под корпусом, вслед за И.А. Пахолковой, мы понимаем «сформированную по определенным правилам выборку данных из области реализации языковой системы» [8. С. 128]. Корпусы текстов являются распространенным материалом для изучения терминов и их совокупностей, поскольку позволяют рассматривать термины в их естественном окружении, открывая широкие возможности для их функционального анализа. При этом лингвистический корпус, по выражению В.А. Плунгяна, – это «не только мощный инструмент исследования языка, но и новая идеология, ориентирующая исследователя на текст как главный объект теоретической рефлексии» [9. С. 14].

Преимущество корпусного метода отбора языковых данных заключается в том, что он «задает объективную формальную процедуру формирования представительной коллекции данных, независимой от произвола исследователя» [10. С. 19]. М.А. Соколова подчеркивает, что крупные корпусы применяются в терминоведении в первую очередь для «установления статистических данных (нормированная частотность ЛЕ на 1 млн словоупотреблений) и для поиска контекстного иллюстративного материала» [11. С. 230]. Для решения более узких исследовательских задач использование крупных корпусов не всегда представляется целесообразным [11. С. 230]. Данное положение можно применить к корпусному изучению терминов лингводидактического стандартизированного тестирования.

Ряд терминов, использующихся в рассматриваемой области, еще не нашли отражения в крупных электронных корпусах немецкого языка. Анализ одного из масштабных корпусов современного письменного немецкого языка DeReKo [12], в частности, его подкорпуса Archiv W как самого крупного архива письменных текстов с XIII в. до настоящего времени [13], показал отсутствие представленности в нем некоторых современных лингводидактических терминов: *Testwirkungsvalidität, textsor-*

tenadäquat, quantitative Validierung, Zielgruppenadäquatheit, plurikulturelle Kompetenz и многих других. Например, термин *Washback*, имеющий в лингводидактике значение «влияние тестирования, экзамена, эвалюации как форм контроля успешности обучения на образовательный процесс» [14. С. 119], встречается в корпусе только один раз как термин пищевой промышленности «бродильный танк» [13]. Лингводидактический термин в форме аббревиатуры *PTN* (*Prüfungsteilnehmende*) также обнаружен только в других, не лингводидактических значениях: *Pilot Training Network, Pata Travel News* и др.

Отсутствие ряда терминов немецкого стандартизированного лингводидактического тестирования в корпусе DeReKo ведет к необходимости составления собственного корпуса. Для этого в исследовании были предприняты следующие шаги. Вначале был отобран и оцифрован практический материал 16 руководств для экзаменаторов. Затем материал был загружен в корпусные менеджеры, при помощи которых происходила дальнейшая работа с терминами, направленная на их отбор из эмпирического материала. К наиболее известным корпусным менеджерам относятся *WordSmith Tools*, *Monoconc Pro*, *AntConc*, *CQP*, *XAIRA*, *NoSketch Engine* [15. С. 23]. В данном исследовании подробнее изучается потенциал применения двух корпусных менеджеров – *Sketch Engine* и *AntConc* – для анализа их возможностей с точки зрения автоматического извлечения специальных лексем из текстов и выявления их коллокаций.

Sketch Engine [16] является одним из ведущих инструментов компьютерной лингвистики, использующихся для лексикографических исследований [17. С. 7], в том числе для решения проблем терминоведения, связанных с извлечением и анализом терминов [17. С. 16]. Сервис разработан компанией *Lexical Computing*, основанной в 2003 г. под руководством британского лингвиста А. Килгарида [18]. *Sketch Engine* обладает широким функционалом, позволяющим автоматически обрабатывать корпус текстов со специальными лексемами из любой предметной области, в частности:

- генерировать список ключевых одно- и многословных терминов (*keywords*);
- формировать список терминов по заданным критериям (*wordlist*) (например, терминов с определенным корнем или единиц заданной части речи);
 - находить N-граммы (N-grams) – последовательности из заданного количества элементов;
 - составлять тезаурус (*thesaurus*) для термина, т.е. список синонимов и слов, «принадлежащих одной категории (семантическому полю)» [19] (здесь и далее перевод наш. – *A.E., A.M.*);
 - показывать конкорданс (*concordance*) заданного термина как «список всех употреблений данного слова в контексте со ссылками на источник» [15. С. 9];

- представлять «эскиз» слова (word sketch), т.е. информацию о сочетаемости данного термина с другими словами, выступающими в качестве различных частей речи (существительных, глаголов, предлогов и др.) в разных синтаксических позициях (в качестве модификаторов, объектов, субъектов и др.);
- сопоставлять пару терминов с точки зрения сходств и различий в их сочетаемости (word sketch difference).

Другой корпусный менеджер – AntConc [20] – также является сервисом, включающим «программные средства для поиска данных в корпусе, получения статистической информации и предоставления результатов пользователю в удобной форме» [15. С. 19]. По сравнению с инструментом Sketch Engine, данная программа, разработанная в 2002 г. британским исследователем Л. Энтони [21. С. 3], является бесплатной, однако имеет более ограниченный функционал, позволяющий составить конкорданс и организовать текстовую информацию для поиска коллокаций, показать частотность слова, выявить список ключевых слов и N-грамм, представить список слов на основе определенной морфемы, вычислить дисперсию слова как меру его распределения в текстах корпуса.

Таким образом, методология исследования включает применение корпусных менеджеров (Sketch Engine, AntConc) в терминоведческом исследовании в сфере стандартизированного лингводидактического тестирования. Выбранные инструменты имеют широкий функционал, позволяющий изучать как отдельно взятые обозначения, так и термины в контекстах. Для оценки потенциала использования корпусных менеджеров с целью анализа немецких лингводидактических терминов в исследовании рассматриваются такие функции программ, как автоматическое извлечение терминов; составление тезауруса и списка терминов по заданным критериям; выявление N-грамм; разработка конкорданса заданного слова; определение списка типичных коллокаций к термину; сопоставление коллокаций двух терминов; вычисление количественных характеристик. Обсуждаются преимущества и недостатки данных функций, выявленные на материале лингводидактических терминов.

Исследование и результаты

Развитие цифровых технологий привело к появлению сервисов с функцией автоматического извлечения терминов, которая также доступна в корпусном менеджере Sketch Engine. В другой программе – AntConc – можно выявить только ключевые слова, среди которых встречаются нетерминологические единицы.

В числе 100 однословных терминов немецкого лингводидактического тестирования, автоматически выделенных в Sketch Engine, программа распознает ряд ключевых терминов: *Prüfungsziel*, *Prüfungsaufgabe*, *Modellsatz*, *Niveaustufe*, *Hörverstehen*. Тем не менее мы признаем

необходимость исключения из списка 13 лексем, поскольку они являются фамилиями ученых (*Perlmann-Balme, Bachman, Glaboniat*) либо дубляжами ранее определенных программой терминов в другой грамматической форме (например, в форме мн. ч. или Р. п.). Е.И. Больщакова и Н.Э. Ефремова также указывают на то, что наличие терминологических вариантов, которые «по возможности должны быть распознаны», является одной из сложностей при автоматическом извлечении терминов из текста [22. С. 120]. Это прослеживается в процессе автоматического извлечения терминов из контекстов. На рис. 1 (№ 2 и 3) показан дубляж форм одного и того же термина *Prüfungsziel* в ед. и мн. ч.; № 20 (*prüfend*) и № 22 (*prüfende*) также пересекаются.

SINGLE-WORDS ✓		MULTI-WORD TERMS ✓	
		reference corpus: German Web 2020 (deTenTen20) (items: 18,558)	
Lemma	Lemma	Lemma	Lemma
1 goethe-zertifikat	14 wirtschaftsdeutsch	27 hörtext	40 sprachliche
2 prüfungsziel	15 modellsatz	28 wortschatz	41 prüfungsmaterialien
3 prüfungsziele	16 prüfungsaufgabe	29 b1	42 rezeptiv
4 testbeschreibung	17 prüfungsteilnehmend	30 b2	43 prüfungsaufgaben
5 prüfungsteilnehmenden	18 sprachhandlung	31 tn	44 rahmencurriculum
6 subtest	19 prüfungsmodule	32 prüfungsformen	45 ptn
7 prüfungsteil	20 prüfend	33 wortliste	46 leitpunkt
8 referenzrahmen	21 textsorte	34 inhaltpunkt	47 ergebnisbogen
9 prüfungsform	22 prüfende	35 akkusativ	48 aufgabenblatt
10 prüfungszentrum	23 niveaustufe	36 monologisch	49 sprachprüfung
11 goethe-institut	24 europarat	37 prüfungsteilnehmer	50 orthografie
12 inventare	25 hörverstehen	38 a2	
13 antwortbogen	26 dativ	39 leseverstehen	

Рис. 1. Фрагмент списка однословных терминов, автоматически извлеченных корпусным менеджером Sketch Engine

Проблемы дублирования словоформ, смешивания терминов и не-терминов, распознания только части многословного термина и другие аспекты обнаружены среди списка лингводидактических терминов, состоящих из двух и более слов. Ряд таких терминов извлечен программой достаточно точно, например: *Gemeinsamer europäischer Referenzrahmen, sprachliches Mittel, sprachliche Handlung, mündliche Interaktion, Niveau A2, kommunikative Aufgabe, Deutsch als Fremdsprache, kommunikative Kompetenz, Durchführung der Prüfung*. При извлечении других терминов программа допустила ошибки: например, представлена только часть термина (*Association of* вместо *Association of Language Testers in Europe, Common European* вместо *Common European Framework of Languages*), в

качестве термина определены фамилия и имя ученого (*Emil Schmank, Michaela Perlmann-Balme*), некоторые однословные термины неправильно соединены в словосочетание (*Arbeitszeit Bewertung* вместо *Arbeitszeit* и *Bewertung*). Помимо этого, отнесение некоторых единиц к списку терминов представляется спорным: часть лексем является не терминами, а номенами, поскольку обозначает единичные объекты (*Aufgabe 2, Teil 3, Subtest 1*). Другая часть лексем не обозначает специальных понятий, в связи с чем не является терминами (*wichtige Einzelheit*). После ручной проверки результатов автоматической выборки многословных лингводидактических терминов 33 единицы были исключены из списка.

Вспомогательным инструментом при выборке терминов и других специальных лексем в Sketch Engine является функция автоматического составления тезауруса (в сервисе AntConc данная функция не представлена). Тезаурус содержит слова, появляющиеся в сходных с ключевым словом контекстах [23]. Например, на основе ключевого слова *Prüfer* программа выстроила тезаурус, состоящий из 50 слов. Некоторые из них являются близкими по значению (*Prüfende, Bewertende, Teilnehmer, Prüferin*), что говорит о достаточно точном определении программой понятийно схожих терминов. На рис. 2 представлены первые 10 слов.

Lempos	Frequency ?
1 Prüfende	129 ...
2 Prüfend	279 ...
3 Bewertende	32 ...
4 Teilnehmer	203 ...
5 Gesprächspartner	167 ...
6 Prüferin	90 ...
7 Schülerin	313 ...
8 Mitschüler	64 ...
9 Trainingsmaterial	30 ...
10 Lehrkraft	72 ...

Рис. 2. Фрагмент тезауруса, автоматически составленный по ключевому слову *Prüfer* в программе Sketch Engine

Однако в сгенерированном тезаурусе также встречаются слова, не обозначающие субъект тестирования, например *Trainingsmaterial* (рис. 2, № 9). Также программа не выявляет термины, обнаруженные в результате ручной выборки: *Assessor, Testentwickler, Prüfungskommision, Testgutachter, Aufsichtsperson, Erstbewertende, Zweitbewertende* и др. Это говорит о необходимости дополнения результатов автоматической обработки текста ручной выборкой исследователей.

Для исследования терминов необходимо выявление их количественных характеристик (частотности употребления). Количественный анализ возможен в корпусном менеджере AntConc благодаря вычислению меры дисперсии (dispersion) термина – статистического показателя

распределения термина в текстах корпуса: чем равномернее слово представлено в разных частях корпуса, тем выше его дисперсия. Например, сравнение дисперсии терминов *Prüfung* и *Test* в сфере лингводидактического стандартизированного тестирования показывает большую распределенность и функциональную нагрузку термина *Prüfung*. Это можно видеть на рис. 3, который показывает распределение термина в анализируемых руководствах (Plot) с конкретными числовыми характеристиками (Dispersion).

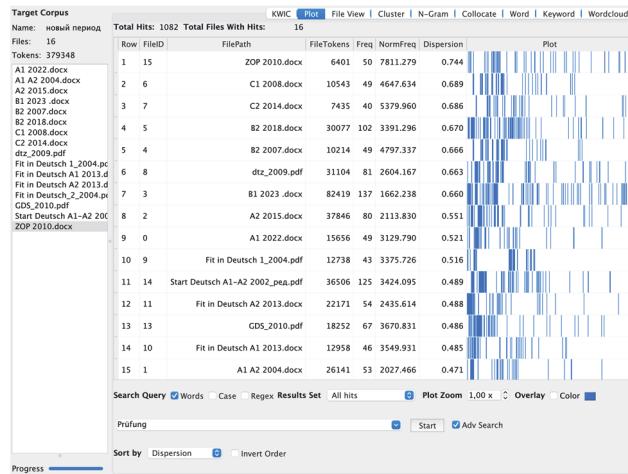


Рис. 3. Дисперсия термина *Prüfung* в корпусном менеджере AntConc (фрагмент)

Дисперсия термина *Test* показана на рис. 4, который наглядно демонстрирует меньшую частотность данного термина по сравнению с термином *Prüfung*.

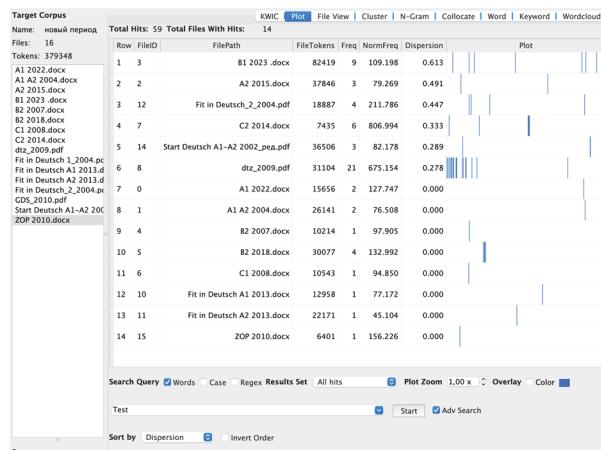


Рис. 4. Дисперсия термина *Test* в корпусном менеджере AntConc

Еще одной функцией при выборке специальных лексем может служить выявление N-грамм, т.е. последовательностей лексем из заданного количества элементов. Sketch Engine автоматически выявляет частотные сочетания единиц от минимум двух до максимум шести компонентов, AntConc ищет N-граммы от одного и более количества элементов. Оба сервиса позволяют задать необходимые критерии поиска (например, сочетания с определенной морфемой). Важно отметить, что в терминоведческом исследовании необходимо ограничивать поиск определенными критериями, в противном случае программы выдадут результаты наиболее частотных сочетаний общелитературного языка (*in der, besteht aus, im Hinblick auf* и т.д.). При поиске N-грамм, содержащих слово *Prüfung*, обе программы выделили терминологические сочетания из двух (*Prüfungsteil Hören, mündliche Prüfung*) и трех (*Prüfung Start Deutsch, Prüfung Goethe-Zertifikat B2*) элементов. Однако наряду с ними сервисы показали нетерминологические сочетания (*Prüfung wird als, Prüfungsziel und, Prüfungsteilnehmenden zeigen* и др.). Поэтому автоматически извлеченные N-граммы требуют дополнительной проверки со стороны специалистов.

Продуктивной функцией корпусных менеджеров Sketch Engine и AntConc является автоматическое составление списка слов по заданным критериям, например по определенной морфеме или слову. Поиск лемм, содержащих морфему *-prüf-*, в программе Sketch Engine выдал 248 словоформ разных частей речи (рис. 5).

Word	Frequency ? ↓	Word	Frequency ? ↓	Word	Frequency ? ↓
1 prüfung	1,802 ...	18 prüfungsteilnehmer	112 ...	35 prüfungszeit	44 ...
2 prüfungsziel	523 ...	19 prüfungszentrum	111 ...	36 prüfungsmaterial	44 ...
3 prüfungsziele	518 ...	20 prüfungsmaterialien	97 ...	37 gruppenprüfung	42 ...
4 prüfungsteilnehmenden	344 ...	21 prüfungsergebnisse	91 ...	38 prüfungsteils	42 ...
5 prüfungsteil	284 ...	22 prüft	78 ...	39 prüfungserstellung	42 ...
6 prüfungen	277 ...	23 prüfungsteile	71 ...	40 prüfern	40 ...
7 prüfungsaufgaben	268 ...	24 prüferinnen	68 ...	41 sprachprüfung	39 ...
8 prüfenden	259 ...	25 prüfungsdurchführung	60 ...	42 prüfungsaufgabe	38 ...
9 prüfungsformen	215 ...	26 paarprüfung	57 ...	43 prüfungsentwicklung	36 ...
10 prüfungsform	211 ...	27 prüfungs	55 ...	44 prüfungsordnung	36 ...
11 prüfende	192 ...	28 sprachprüfungen	55 ...	45 prüfungsgebühr	34 ...
12 prüfungsteilnehmende	161 ...	29 prüfungszentrale	54 ...	46 prüfungssatz	33 ...
13 prüfungsmodule	156 ...	30 prüfungsteilen	52 ...	47 überprüfen	33 ...
14 prüfer	153 ...	31 prüfungsablauf	52 ...	48 prüfungsbedingungen	29 ...
15 geprüft	138 ...	32 prüfungsunterlagen	51 ...	49 prüfungsarbeiten	28 ...
16 prüfungszentren	138 ...	33 oberstufenprüfung	50 ...	50 deutschprüfung	28 ...
17 überprüft	132 ...	34 einzelprüfung	44 ...		

Рис. 5. Фрагмент результата поиска лемм, содержащих морфему *-prüf-*, в программе Sketch Engine

Как показывают данные, с помощью сервиса выявлены существительные (*Prüfung, Prüfungsziel, Prüfende, Prüfungsentwicklung, Fit-in-*

Deutsch-1-Prüfung), глаголы (*prüfen, überprüfen, abprüfen*), прилагательные и причастия (*prüfungsrelevant, prüfend, geprüft*).

При запросе о выявлении глаголов, содержащих морфему *-wert-*, программа перечислила необходимые глаголы в порядке уменьшения частотности: *bewerten* (236 употреблений), *auswerten* (36 употреблений), *werten* (13 употреблений), *mitbewerten* (два употребления) (рис. 6).

Verb	Frequency ↓
1 bewerten	236 ***
2 auswerten	36 ***
3 werten	13 ***
4 mitbewerten	2 ***

Рис. 6. Фрагмент результата поиска глаголов, содержащих морфему *-wert-*, в программе Sketch Engine

При этом сервис посчитал не только глаголы в форме инфинитива (№ 45), но и другие словоформы: Partizip II (№ 41, 43, 44), Präteritum (№ 42). В качестве примера можно привести контексты на основе глагола *bewerten* на рис. 7.

<input type="checkbox"/> Details	Left context	KWIC	Right context
41 <input type="checkbox"/> ⓘ doc#3 gt.</></>Nach den ersten Aufnahmen wurden sieben Leistungsbeispiele bewerten , kommentiert und die Bewertungen miteinander verglichen Ziel war eine V			
42 <input type="checkbox"/> ⓘ doc#3 gefasst sowie die Zahl der Folienabbildungen reduziert.</></>Die Partner bewerteten - im Rahmen der Erstellung von zweiten Filmaufnahmen für die Trainingsr			
43 <input type="checkbox"/> ⓘ doc#3 wurden noch einmal die Aufgabe 1 sowie zehn weitere Leistungsbeispiele bewertet , jeweils wieder zunächst in Einzelarbeit und nach Darlegung der Ergebnis			
44 <input type="checkbox"/> ⓘ doc#3 didatenleistungen nach Vertrautmachen mit der Aufgabenstellung.</></> Bewertet wurden insgesamt 22 Kandidatenleistungen, davon jeweils 1 für die Produt			
45 <input type="checkbox"/> ⓘ doc#3 ng Meinungen, Ansichten ausdrücken Beurteilung von Handlungen positiv bewerten dankend anerkennen bedauern Rechtfertigung begründen, rechtfertigen Bi			

Рис. 7. Фрагмент контекстов, содержащих ключевое слово *bewerten*, в программе Sketch Engine

Продуктивной функцией для выявления терминологических сочетаний и коллокаций является составление конкорданса заданного слова, что возможно в обоих корпусных менеджерах. Сервисы позволяют организовать текстовую информацию в зависимости от позиционного типа контекста – правого и левого. Благодаря этому исследователь может выявить типичные коллокации для единицы. Например, анализ правого контекста термина *Fehler* позволяет выявить такие коллокации, как *Fehler in Syntax, Fehler in Morphologie, Fehler in der Rechtschreibung, Fehler in der Satzstellung* (рис. 8).

Левый контекст термина *Fehler* в корпусном менеджере AntConc представлен на рис. 9 такими терминами, как *morphologische Fehler, semantische Fehler, syntaktische Fehler, Syntaxfehler, sprachliche Fehler, Wortschatzfehler, nicht-systematische Fehler*, и др.

15	Fit in ...	erschweren. An mehreren Stellen beeinträchtigen die	Fehler	in Syntax, Morphologie (und Orthografie/Interpunktio
16	Fit in ...	erschweren. An mehreren Stellen beeinträchtigen die	Fehler	in Syntax, Morphologie (und Orthografie/Interpunktio
17	Start ...	edergegeben; Textsorte halbe Punktzahl, da zu viele	Fehler	in Anrede und Gruß. 1 72 4 PRÜFUNGSFORMEN STAR
18	Fit in ...	unkte). Das entspricht 20 % des Gesamtergebnisses.	Fehler	in der Orthografie werden nur dann mit Punktabzug I
19	Fit in ...	ntke). Das entspricht 20 % des Gesamt- ergebnisses.	Fehler	in der Orthografie werden nur dann mit Punktabzug I
20	Fit in ...	n Familienname ist Schmidt. r Fehler, - Ich hatte vier	Fehler	in der Englischarbeit: s Fenster, - Ich möchte einen P
21	Fit in ...	tur faulenzen. Hier fehlt ein Buchstabe. Ich hatte vier	Fehler	in der Englischarbeit: Ich habe meinen Geburtstag ge
22	A1 A2 ...	Kommentar Aufgabe voll erfüllt, Inhalt verständlich.	Fehler	in der Grusformel unerheblich. Textsorte: Kein Punkt
23	ZOP ...	er inhaltlich eindeutig verständlich und sinnvoll sind.	Fehler	in der Rechtschreibung und Grammatik werden bei d
24	B2 ...	iben Sie die richtige Form an den Rand. (Beispiel 01)	Fehler	in der Satzstellung: Schreiben Sie das falsch platziert
25	B2 ...	einem kurzen Artikel äußern, in Aufgabe 2 sollen sie	Fehler	in einem Brief korrigieren. Die beiden Textvorlagen h

Рис. 8. Правый контекст термина *Fehler* в корпусном менеджере AntConc (фрагмент)

140	B2 ...	ir syntaktische, morphologische und semantische	Fehler	enthält. Es müssen zehn Fehler gefunden und kor
141	B2 ...	n in der Lage sind, syntaktische und semantische	Fehler	in einem vorgegebenen Brief zu erkennen und zu
142	Start ...	j/oder Aufgabe noch erfüllt, mehrere sprachliche	Fehler,	Frage / Bitte nicht erfasst und/oder Aufgabe nicht
143	Start ...	j/oder Aufgabe noch erfüllt, mehrere sprachliche	Fehler,	Frage / Bitte nicht eifasst und/ oder Aufgabe nich
144	Start ...	mer verständlich, gute Syntax nur kleine Syntax-,	Wortschatzfehler 2	Gesamt Kommentar F A 0,5 0,5 2,5 Frage und Ant
145	Start ...	erfüllt, immer verständlich, gute Syntax Syntax-,	Wortschatzfehler	Teil 2: Ein Alltags- gespräch führen 1 F A 0,5 PTN
146	B2 ...	eigentliche Ausrutscher oder nicht-systematische	Fehler	und kleinere Mängel im Satzbau können vorkommr
147	B2 ...	eigentliche Ausrutscher oder nicht- systematische	Fehler	und kleinere Mängel im Satzbau können vorkommr

Рис. 9. Левый контекст термина *Fehler* в корпусном менеджере AntConc (фрагмент)

Оба корпусных менеджера позволяют автоматически определить список типичных коллокаций к термину. Функция Collocates в AntConc дает возможность сделать «мгновенный статистический анализ по словам, стоящим слева и справа от искомого элемента» [24. С. 41]. Например, среди 10 основных коллокатов к термину *Prüfung* указаны слова *die, der, mündliche, bestanden, deutsch* и др. Sketch Engine не только определяет коллокации слова, но и группирует их по различным критериям, формируя «эскиз слова» (word sketch). Корпусный менеджер достаточно четко определяет эскиз термина *Prüfung*: он сочетается с рядом атрибутов (*mündliche, gesamte, bestandene, schriftliche, einundhalbstündige, elektronische Prüfung*), глаголов (*bestehen, beziehen, dauern, basieren, bieten, erbringen, verwenden, überprüfen, orientieren*), объектов в Р. п. (*Prüfungen des Goethe-Instituts*) и др. (рис. 10).

В некоторых контекстах программа допустила ошибки: например, выделила только глагол *finden* вместо глагола с отделяемой приставкой *stattfinden* (*die Prüfung findet nach Bedarf statt*), а также неверно установила ряд объектов в Р. п., сочетающихся с термином *Prüfung* (*Prüfungserstellung, Prüfung, Version*). Это говорит о необходимости проверки специалистами результатов автоматического анализа.

Другая функция Sketch Engine, применимая в терминоведческом исследовании, связана с сопоставлением коллокаций двух терминов (word sketch difference), что может помочь при выявлении семантических отношений между ними. Например, сопоставление терминов *Stufe*

и *Niveau* показало, что обе единицы имеют сходное окружение – сочетаются с лексемами, связанными с уровневым владением языком (*die zweite Stufe, A2* и *Niveau A2*), пример 1:

(1) *Für Schülerinnen und Schüler, die eine Fremdsprache lernen, steht die zweite Stufe, A2, für eine beachtliche Leistung. Die Existenz der ersten international bekannten Deutschprüfung auf dem Niveau A2 Fit in Deutsch 2 ist einer Initiative des italienischen Unterrichtsministeriums zu verdanken* [6. S. 61].

modifiers of "Prüfung"	verbs with "Prüfung" as subject	genitive objects of "Prüfung"
mündlich der mündlichen Prüfung	bestehen Die Prüfung besteht aus drei Teilen	Goethe-Institut Prüfungen des Goethe-Instituts
gesamt den Meßfehler der gesamten Prüfung zu ermitteln	beziehen dass ihre Prüfung sich auf ein theorethisches Konstrukt bezieht , z. B.	Aufgabenformate Inhalte und der Aufgabenformate einen sorgfältigen Abgleich der Prüfung mit den Kannbeschreibungen
bestanden Bei nicht bestandener Prüfung erhält	dauern Die Prüfung dauert circa	Prüfungserrstellung Datenschutz Arbeitsschritte der Prüfungserrstellung Die Prüfungen des Goethe-Instituts werden
schriftlich Die schriftliche Prüfung	erbringen den die Prüfung erbringt , und über	Prüfung der Prüfung Die Prüfung
unmittelbar Unmittelbar vor der mündlichen Prüfung erhalten die Teilnehmenden	bieten berücksichtigt . Die Prüfung bietet	Version Kennzeichen der beiden Versionen der Prüfung für Jugendliche und
zweiteilig In der zweiteiligen Prüfung sollen die Prüfungsteilnehmenden	basieren Grundlage Die Prüfung basiert auf dem in	
zentral Erstellung und Durchführung zentraler Prüfungen des Goethe-Instituts und	verwenden durchgeführt . Die Prüfung verwendet gedruckte Prüfungsunterlagen	
eineinhalbstündig sich um eine eineinhalbständige Prüfung , die	überprüfen Prüfung fit in Deutsch 1 überprüft	
neu der neuen Prüfung	finden Prüfungsgerümine Die Prüfung findet nach Bedarf statt	
elektronisch wird nicht als elektronische Prüfung durchgeführt . Das	orientieren Die Prüfung orientiert sich	
	greifen	

Рис. 10. Фрагмент «эскиза» термина *Prüfung* в корпусном менеджере Sketch Engine

Термины *Stufe* и *Niveau* также входят в коллокации с обозначениями Общеевропейских компетенций владения иностранным языком, что показано на примерах 2 и 3:

(2) *Die Lernenden sollen Deutschkenntnisse bereits auf der ersten Stufe des Gemeinsamen europäischen Referenzrahmens für Sprachen¹ dokumentieren können* [7. S. 10].

(3) *Die Anpassung an die Niveaus des Referenzrahmens geschieht auf der methodischen Basis des Manuals for Relating Language Examinations to the Common European Framework of Reference for Languages des Europarates (Council of Europe 2009)* [5. S. 21].

В сочетаемости данных терминов также есть различия. Согласно анализу в программе Sketch Engine, термин *Stufe* имеет тенденцию к использованию в сочетаниях с терминами *Kompetenzskala* (пример 4):

(4) *Sie setzt Sprachkenntnisse auf der untersten Stufe – A1 – der sechsstufigen Kompetenzskala des Gemeinsamen europäischen Referenzrahmens voraus* [7. S. 8].

В то же время термин *Niveau* имеет тенденцию к образованию коллокаций с лексемой *Prüfung* (пример 5):

(5) *Wortschatz und Syntax entsprechen dabei dem Niveau der Prüfung* [7. S. 51].

Таким образом, можно получить данные, оптимизирующие терминоведческое исследование, при этом нужно учитывать, что интерпретация данных зависит от исследователя.

Заключение

Процесс выборки терминов является одним из центральных в терминоведческой работе. Специфика немецких терминов стандартизированного лингводидактического тестирования состоит в их консубстанциональности, т.е. формальном совпадении с общеупотребительными словами. Терминологическое значение консубстанциональных лингводидактических терминов проявляется в определенном контексте, поэтому при их выборке из специальных текстов необходимо обращаться к коллокациям. Оптимизация процесса отбора лингводидактических терминов, в том числе с учетом их коллокаций, возможна благодаря применению корпусных менеджеров.

Обращение к корпусным менеджерам Sketch Engine и AntConc в терминоведческом исследовании позволяет оптимизировать процесс выборки терминов при помощи составления собственного корпуса специальных контекстов, представления терминологической информации в структурированном виде и ее автоматической обработки. В зависимости от использования того или иного корпусного менеджера, исследователь получает доступ к широкому набору функций, включающих автоматическое извлечение ключевых слов и терминов (одно- и многословных), составление тезауруса, вычисление меры дисперсии термина, выявление N-грамм, создание списка слов по заданным критериям (определенной морфеме, части речи), разработку конкорданса и списка типичных коллокаций, а также сопоставление коллокаций двух терминов. Как показывает исследование, корпусные менеджеры, обладая рядом технических преимуществ, допускают ошибки, связанные с неправильной либо не полной выборкой терминов. В связи с этим специалистам необходимо вручную корректировать, дополнять и интерпретировать результаты, полученные в результате автоматической обработки текстов.

Список источников

1. **Комарова З.И.** Информационные вызовы современного общества: лингвотерминоведческие и терминографические процессы // Политическая лингвистика. 2017. № 5 (65). С. 12–34.

2. **Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С.** Компьютерная лингвистика: задачи, подходы, ресурсы // Автоматическая обработка текстов на естественном языке и анализ данных. М. : Изд-во НИУ ВШЭ, 2017. С. 7–30.
3. **Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В.** Часть I. Основы теоретической, вычислительной и экспериментальной лингвистики, или Размышления о месте лингвиста в компьютерной лингвистике // Автоматическая обработка текстов на естественном языке и компьютерная лингвистика. М. : МИЭМ, 2011. С. 7–89.
4. **Комарова З.И.** Методология, метод, методика и технология научных исследований в лингвистике : учеб. пособие. Екатеринбург : Изд-во УрФУ, 2012. 818 с.
5. **Glaboniat M., Perlmann-Balme M., Studer T.** Zertifikat B1. Prüfungsziele. Testbeschreibung. München : Hueber, 2018. 230 S.
6. **Hennemann D., Karamichali E., Perlmann-Balme M., Stelter C.** Goethe Zertifikat A2. Prüfungsziele. Testbeschreibung. München : Hueber, 2015. 117 S.
7. **Perlmann-Balme M., Stoffers M.** Goethe Zertifikat A1: Start Deutsch 1. Prüfungsziele. Testbeschreibung. München : Goethe Institut, 2022. 107 S.
8. **Пахолкова И.А.** Применение методов корпусной лингвистики в традиционном языкоznании // Вестник МГЛУ. 2012. Вып. 13 (646). С. 125–135.
9. **Плунгян В.А.** Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики // Русский язык в научном освещении. 2008. № 2 (16). С. 7–20.
10. **Кибрек А.Е., Брыкина М.М., Леонтьев А.П., Хитров А.Н.** Русские посессивные конструкции в свете корпусно-статистического исследования // Вопросы языкоznания. 2006. № 1. С. 16–45.
11. **Соколова М.А.** Использование корпусных инструментов в исследовании терминологии (на материале английского политического языка) // Политическая лингвистика. 2024. № 4 (106). С. 229–234.
12. **Das Deutsche Referenzkorpus – DeReKo.** URL: <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/> (дата обращения: 15.12.2024).
13. **Archive** // IDS. Leibniz-Institut für deutsche Sprache. URL: <https://www2.ids-mannheim.de/cosmas2/projekt/referenz/archive.html> (дата обращения: 15.12.2024).
14. **Едличко А.И., Захарова О.Л., Мишинова А.В.** Терминологический справочник к программе «Учимся обучать немецкому» = Fachlexikon zu DLL – Deutsch Lehren Lernen / под ред. А.И. Едличко, О.Л. Захаровой. М. : Изд-во Моск. ун-та, 2022. 191 с.
15. **Моделирование** в корпусной лингвистике: специализированные корпусы русского языка / В.П. Захаров, И.В. Азарова, О.А. Митрофанова, А.М. Попов, М.В. Хохлова ; отв. ред. В.П. Захаров. СПб. : Изд-во СПбГУ, 2019. 208 с.
16. **Sketch Engine.** URL: <https://www.sketchengine.eu/> (дата обращения: 20.12.2024).
17. **Kilgarriff A., Baisa V., Bušta J., Jakubíček M., Kovář V., Michelfeit J. et al.** The Sketch Engine: ten years on. Lexicography // Lexicography ASIALEX. 2014. Vol. 1. P. 7–36.
18. **Lexical Computing** // Lexical Computing. URL: <https://www.lexicalcomputing.com/lexical-computing/> (дата обращения: 07.12.2024).
19. **Thesaurus** – synonyms, antonyms and similar words // Sketch Engine. URL: <https://www.sketchengine.eu/guide/thesaurus-synonyms-antonyms-similar-words/> (дата обращения: 7.12.2024).
20. **AntConc** // Laurence Anthony's Website. URL: <https://www.laurenceanthony.net/software/antconc/> (дата обращения: 12.12.2024).
21. **Introduction** to AntConc // NLS Foundry. URL: <https://nlsfoundry.s3.amazonaws.com/reports/nls-reports-intro-antconc-march-2023.pdf> (дата обращения: 04.12.2024).
22. **Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С.** Извлечение информации из текстов: портрет

направления // Автоматическая обработка текстов на естественном языке и анализ данных. М. : Изд-во НИУ ВШЭ, 2017. С. 83–126.

23. **Distributional** thesaurus // Sketch Engine. URL: <https://www.sketchengine.eu/glossary/distributional-thesaurus/> (дата обращения: 13.01.2025).
24. **Комарова И.А.** Корпусные исследования с помощью сервиса AntConc в условиях работы в вузе // Язык и культура. 2020. № 52. С. 36–50.

References

1. Komarova Z.I. (2017) Informatsionnye vyzovy sovremennoego obshchestva: lingvoterminovedcheskie i terminograficheskie protsessy [Information challenges of modern society: linguoterminological and terminographic processes] // Politicheskaya lingvistika. 5 (65). pp. 12–34.
2. Bol'shakova E.I. (2017) Komp'yuternaya lingvistika: zadachi, podkhody, resursy [Computational linguistics: tasks, approaches, resources] // Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i analiz dannykh / Bol'shakova E.I., Vorontsov K.V., Efremova N.E., Klyshinskij E.S., Lukashevich N.V., Sapin A.S. M.: Izdatel'stvo NIU VShE. pp. 7–30.
3. Yagunova E.V. (2011) Chast' I. Osnovy teoretycheskoj, vychislitel'noj i eksperimental'noj lingvistiki, ili razmyshleniya o meste lingvista v komp'yuternoj lingvistike [Part I. Basics of theoretical, computational, and experimental linguistics, or reflections on the linguist's place in computational linguistics] // Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i komp'yuternaya lingvistika / Bol'shakova E.I., Klyshinskij E.S., Lande D.V., Noskov A.A., Peskova O.V., Yagunova E.V. M.: MIEM. pp. 7–89.
4. Komarova Z.I. (2012) Metodologiya, metod, metodika i tekhnologiya nauchnykh issledovanij v lingvistike: uchebnoe posobie [Methodology, method, and techniques of linguistic research: a textbook]. Ekaterinburg: Izdatel'stvo UrFU. 818 p.
5. Glaboniat M., Perlmann-Balme M., Studer T. (2018) Zertifikat B1. Prüfungsziele. Testbeschreibung. München: Hueber. 230 p.
6. Hennemann D., Karamichali E., Perlmann-Balme M., Stelter C. (2015) Goethe Zertifikat A2. Prüfungsziele. Testbeschreibung. München: Hueber. 117 p.
7. Perlmann-Balme M., Stoffers M. (2022) Goethe Zertifikat A1: Start Deutsch 1. Prüfungsziele. Testbeschreibung. München: Goethe Institut. 107 p.
8. Pakholkova I.A. (2012) Primenenie metodov korpusnoj lingvistiki v traditsionnom yazykoznanii [Application of corpus linguistics methods in traditional linguistics] // Vestnik MGLU. 13 (646). pp. 125–135.
9. Plungyan V.A. (2008) Korpus kak instrument i kak ideologiya: o nekotorykh urokakh sovremennoj korpusnoj lingvistiki [Corpus as a tool and as ideology: lessons from modern corpus linguistics] // Russkij jazyk v nauchnom osveshchenii. 2 (16). pp. 7–20.
10. Kibrik A.E., Brykina M.M., Leont'ev A.P., Khitrov A.N. (2006) Russkie possessivnye konstruktsii v svete korpusno-statisticheskogo issledovaniya [Russian possessive constructions in light of corpus-statistical research] // Voprosy yazykoznanija. 1. pp. 16–45.
11. Sokolova M.A. (2024) Ispol'zovanie korpusnykh instrumentov v issledovanii terminologii (na materiale anglijskogo politicheskogo jazyka) [Using corpus tools in terminology research (based on English political language)] // Politicheskaya lingvistika. 4 (106). pp. 229–234.
12. Das Deutsche Referenzkorpus – DeReKo. URL: <https://www.ids-mannheim.de/digspra/kl/projekte/korpora/> (Accessed: 15.12.2024).
13. Archive // IDS. Leibniz-Institut für deutsche Sprache. URL: <https://www2.ids-mannheim.de/cosmas2/projekt/referenz/archive.html> (Accessed: 15.12.2024).
14. Edlichko A.I., Zakharova O.L., Mishnova A.V. (2022) Terminologicheskij spravochnik k programme ‘Uchimsya obuchat’ nemetskomu’ = Fachlexikon zu DLL – Deutsch Lehren

- Lernen [Terminological dictionary for the program “Learning to teach German”] / ed. by A.I. Edlichko, O.L. Zakharova. M.: Izdatel'stvo Moskovskogo universiteta. 191 p.
15. Modelirovaniye v korpusnoj lingvistike: spetsializirovannye korpusy russkogo yazyka (2019) [Modeling in corpus linguistics: specialized corpora of the Russian language] / V.P. Zakharov, I.V. Azarova, O.A. Mitrofanova, A.M. Popov, M.V. Khokhlova; ed. by V.P. Zakharov. SPb.: Izdatel'stvo SPbGU. 208 p.
16. Sketch Engine. URL: <https://www.sketchengine.eu/> (Accessed: 20.12.2024).
17. Kilgarriff A., Baisa V., Bušta J., Jakubíček M., Kovář V., Michelfeit J. et al. (2014) The Sketch Engine: ten years on. Lexicography. Lexicography ASIALEX. Vol. 1. pp. 7–36.
18. Lexical Computing. Lexical Computing. URL: <https://www.lexicalcomputing.com/lexical-computing/> (Accessed: 07.12.2024).
19. Thesaurus – synonyms, antonyms and similar words. Sketch Engine. URL: <https://www.sketchengine.eu/guide/thesaurus-synonyms-antonyms-similar-words/> (Accessed: 7.12.2024).
20. AntConc. Laurence Anthony's Website. URL: <https://www.laurenceanthony.net/software/antconc/> (Accessed: 12.12.2024).
21. Introduction to AntConc. NLS Foundry. URL: <https://nlsfoundry.s3.amazonaws.com/reports/nls-reports-intro-antconc-march-2023.pdf> (Accessed: 04.12.2024).
22. Bol'shakova E.I., Efremova N.E. (2017) Izvlechenie informatsii iz tekstov: portret napravleniya [Information extraction from texts: overview of the field] // Avtomaticheskaya obrabotka tekstov na estestvennom yazyke i analiz dannykh / Bol'shakova E.I., Vorontsov K.V., Efremova N.E., Klyshinskij E.S., Lukashevich N.V., Sapin A.S. M.: Izdatel'stvo NIU VShE. pp. 83–126.
23. Distributional thesaurus. Sketch Engine. URL: <https://www.sketchengine.eu/glossary/distributional-thesaurus/> (Accessed: 13.01.2025).
24. Kotyurova I.A. (2020) Korpusnye issledovaniya s pomoshch'yu servisa AntConc v usloviyah raboty v vuze [Corpus studies using AntConc in university teaching] // Yazyk i kul'tura. 52. pp. 36–50.

Информация об авторах:

Едличко А.И. – кандидат филологических наук, доцент, и.о. заведующей кафедрой немецкого языка и культуры, Московский государственный университет им. М.В. Ломоносова (Москва, Россия). E-mail: ang299@yandex.ru

Мишинова А.В. – преподаватель, Московский государственный университет им. М.В. Ломоносова (Москва, Россия). E-mail: mishnovaa@mail.ru

Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Edlichko A.E., Ph.D. (Philology), Associate Professor, Acting Chair of the Department of the German Language and Culture, Lomonosov Moscow State University (Moscow, Russia). E-mail: ang299@yandex.ru

Mishnova A.V., lecturer of the Chair of German Language and Culture, Lomonosov Moscow State University (Moscow, Russia). E-mail: mishnovaa@mail.ru

The authors declare no conflicts of interests.

Поступила в редакцию 02.09.2025; принята к публикации 28.11.2025

Received 02.09.2025; accepted for publication 28.11.2025