



Информация для цитирования:

Рябцева Н. К. Межъязыковые соответствия и актуальные проблемы терминографии / Н. К. Рябцева // Научный диалог. — 2023. — Т. 12. — № 8. — С. 211—227. — DOI: 10.24224/2227-1295-2023-12-8-211-227.

Riabtseva, N. K. (2023). Cross-Linguistic Correspondences and Current Problems in Terminography. *Nauchnyi dialog*, 12 (8): 211-227. DOI: 10.24224/2227-1295-2023-12-8-211-227. (In Russ.).



Журнал включен в Перечень ВАК

DOI: 10.24224/2227-1295-2023-12-8-211-227

**Межъязыковые
соответствия
и актуальные проблемы
терминографии**

Рябцева Надежда Константиновна

orcid.org/0000-0002-2042-4615

ResearcherID S-7138-2016

Scopus Author ID 55218430500

доктор филологических наук,
ведущий научный сотрудник
отдела прикладного
языкознания

nadia_riabceva@mail.ru

Институт языкознания
Российской академии наук
(Москва, Россия)

**Cross-Linguistic
Correspondences
and Current Problems
in Terminography**

Nadezhda K. Riabtseva

orcid.org/0000-0002-2042-4615

ResearcherID S-7138-2016

Scopus Author ID 55218430500

Doctor of Philology,
leading research scientist,
Department of Applied Linguistics
nadia_riabceva@mail.ru

Institute of Linguistics
of the Russian Academy of Sciences
(Moscow, Russia)

ОРИГИНАЛЬНЫЕ СТАТЬИ

Аннотация:

В статье показано, что в свете цифровых технологий современная лексикографическая и терминологическая деятельность включает множество разнообразных аспектов и типов работы и поэтому превращается в автономное и ведущее лингвистическое направление. Особо актуальную проблему при этом занимает «межязыковая терминография» — составление, в частности, двуязычных словарей и цифровых ресурсов по переводу исходно аутентичной англоязычной многокомпонентной и лингвоспецифичной терминологии на родной, в том числе русский, язык. На материале больших корпусных данных, собранных за последние пять лет (свыше 1 млн словоупотреблений) и включающих как ведущие англоязычные журналы и книги по смежным с цифровыми технологиями лингвистическим дисциплинам (Natural Language Processing — NLP), так и соответствующие им русскоязычные издания, содержащие переводы названий, аннотаций и ключевых слов с русского языка на английский, а также на основе анализа работы автоматических систем перевода установлено, что современные англо-русские терминологические словари и системы по автоматической обработке естественного языка нуждаются в принципиальном совершенствовании. Ведущую роль в нем занимает использование лингвистических знаний о межязыковых соответствиях, в том числе относительно «переводческих универсалий», таких, как экспликация, импликация и нормализация.

Ключевые слова:

межязыковая терминография; терминологические модели; лингвоспецифичность; переводческие универсалии.

ORIGINAL ARTICLES

Abstract:

The paper investigates the current development of digital technologies, its terminology, the problems in cross-linguistic terminography and shows that at present terminography becomes an autonomous and leading multi-dimensional linguistic digital discipline. Its most urgent task is to create effective digital terminological resources of cross-linguistic correspondences to authentic extended language-specific and idiomatic digital terminology in English, including its equivalents in Russian. A wide collection of corpus data, compiled for the last five years (over 1 mil. words) out of leading authentic journals and books in English on Natural Language Processing (NLP) and corresponding authentic texts in Russian, provided with translation of titles, abstracts and key words from Russian into English, as well as the analysis of computer translation systems, shows that contemporary English-Russian dictionaries and digital lexicographic resources need serious improvement. It can be based on contemporary linguistic knowledge in cross-linguistic English–Russian and Russian–English correspondences, particularly concerning “translation universals”, such as explication, implication and normalization.

Key words:

cross-linguistic terminography; language-specific terminology patterns; translation universals.

Межъязыковые соответствия и актуальные проблемы терминографии

© Рябцева Н. К., 2023

1. Введение = Introduction

Современный научный и технический язык, и особенно терминология, постоянно изменяются, пополняются новыми понятиями и терминами, развиваются и даже приобретают новые формы и виды. При этом современные цифровые технологии не только помогают создавать новые типы и виды словарей, в том числе и терминологических, но и сами порождают беспрецедентное количество новых терминов, новые многокомпонентные модели их построения, новые способы их передачи на иностранный (в том числе русский) язык и т. д. Причем «цифровая» терминология вызывает самые большие проблемы у электронных переводчиков, которые еще не отказались от дословного перевода.

Так, в известной статье «The Future of Dictionaries, Dictionaries of the Future» [Nielsen, 2013] подчеркивается, что в настоящее время лексикография как самостоятельная лингвистическая деятельность по своей сути может быть отнесена к типу «a branch of reference science» или даже к «information science». Причем лексикографическая и терминологическая деятельность включает такие разнообразные виды работы, как разработка истории словарей, их типология, составление, концептуальное обоснование структуры и содержания, а также удовлетворение потребностей в образовании, особенно в профессиональном, развитие e-lexicography и т. д., и поэтому она уже давно превращается в автономное и ведущее лингвистическое направление. Цифровые технологии также ставят перед современной лексикографией всех типов и видов, особенно терминологической, принципиально новые проблемы: использование корпусных данных и других цифровых инструментов (например, ‘word sketches’ [Kilgariff, 2006] и мн. др.), придание лексикографическим продуктам цифровой формы — превращение их в e-lexicography, превращение обычных словарей в «цифровых ассистентов» для порождения письма и речи на родном и иностранном языке [Nielsen, 2013], совмещение словарей общеупотребительной и терминологической лексики и их единообразное устройство ([Fuertes-Olivera et al., 2020, с. 91, 108; Tarp, 2022, с. 66] и мн. др.). Кроме

того, современная лексикография и терминография нуждаются в новом, «пользовательском» взгляде на их устройство (a user-oriented approach in making dictionaries). При этом особое внимание в настоящее время уделяется лексикографическим и терминологическим средствам, отражающим динамический характер терминологии, которая постоянно изменяется, обновляется и совершенствуется [Fuertes-Olivera, 2022, с. 125—132]. С этой целью, в частности, определяются основные недостатки интернет-словарей и предлагаются способы их устранения.

Так что в свете современных цифровых технологий лексикографическая и терминологическая деятельность включает множество разнообразных аспектов и типов работы и поэтому превращается в автономное, притом ведущее лингвистическое направление. Особую актуальность имеет «межъязыковая терминография» — составление, в частности, двуязычных словарей по переводу исходно аутентичной англоязычной многокомпонентной и лингвоспецифичной терминологии на родной, в том числе русский, язык.

В данной статье используется мировой опыт в области установления и квалификации межъязыковых соответствий, а также демонстрируются разработанные автором принципы установления англо-русских и русско-английских межъязыковых соответствий, их включения в лексикографические и терминологические источники, а также особенности установления межъязыковых терминологических соответствий между русским и английским языками. Они могут быть использованы в разнообразных прикладных и теоретических областях — от преподавания английского языка и перевода с английского языка на русский и обратно до составления качественных лексикографических и терминологических ресурсов и интерпретации эволюционных изменений в грамматических и терминологических системах соответствующих языков.

Статья состоит из введения и трех разделов. В разделе 1 описаны исходные положения (продолжающегося) исследования. В разделе 2 указываются материал и методы исследования, а также раскрывается суть актуальных проблем в установлении межъязыковых соответствий. В разделе 3 показывается эффективный теоретический аппарат решения межъязыковых лексикографических и терминологических соответствий: рассматриваются универсальные переводческие приемы перевода, в том числе и перевода (новой) терминологии, признанные мировым сообществом. Это — экспликация, импликация и нормализация. Демонстрируются также авторские приемы перевода развернутых терминов в области инновационной цифровой обработки естественного языка (Natural Language Processing — NLP). В заключение указываются перспективы дальнейших исследований.

2. Материал, методы, обзор = Material, Methods, Review

В настоящее время словарное дело из «производства и издания словарей», особенно терминологических, превращается в организацию специальных лексикографических и терминологических информационных служб, которые приобретают совершенно новые черты: ориентация на определенного пользователя и потому на информационное общество, основанное на знании, и мн. др. При этом они превращаются в информационные инструменты, которые обеспечивают индивидуальную поддержку различным, особенно профессиональным, группам пользователей, нуждающимся в лексикографически (и терминологически) значимой, направленной именно им информации (*lexicographically and terminologically relevant data*), и производят ее «персонализацию» и анализ динамического взаимодействия на поверхностном и глубинном уровне. Это не значит, что уже завтра у нас появятся принципиально новые цифровые лексикографические продукты и терминологические службы, но уже сегодня для этого много делается, особенно специалистами, анализирующими и совершенствующими работу действующих лексикографических и терминологических ресурсов, большинство из которых нуждается в существенной доработке [Huang et al., 2021, с. 77].

В данной статье на материале больших корпусных данных (свыше 1 млн словоупотреблений), собранных (автором) за последние пять лет и включающих как ведущие англоязычные журналы и книги по смежным с цифровыми технологиями лингвистическим дисциплинам, так и русскоязычные издания (книги, журналы, материалы конференций и др.), содержащие переводы названий, аннотаций и ключевых слов с русского языка на английский, а также на основе анализа работы автоматических систем перевода (*Google Translate, Yandex Translate, DeepL, ABBY* и др.) установлено, что современные англо-русские терминологические словари и системы по автоматической обработке естественного языка нуждаются в принципиально новом подходе, ведущую роль в котором занимает использование лингвистических знаний о межязыковых соответствиях. Причем ориентация на внедрение нейросетей в соответствующие процедуры на основе самообучения не обязательно покажет желаемые результаты уже хотя бы потому, что необходимых для выполнения этой задачи текстов, самых современных двужычных, параллельных, пока не существует (ср. [Мифтахова и др., 2019]).

Особое внимание при этом в современной лингвистике уделяется одной из наиболее остро стоящих терминологических проблем — переводу на родной язык и с родного языка многокомпонентных терминов, которые все более активно появляются во всех отраслях знаний и составляют в них более 75 % от всей терминологии, см.: [Пегов, 2018; Пегов, 2019; Каменева, 2019; Станкевич, 2020 и мн. др.]. Назовем их важнейшие черты: появление

исходно в английском языке, их увеличивающаяся точность и «прецизионность» номинации, структурная усложненность и потому в большинстве случаев невозможность их пословного перевода, который используется в современных автоматических переводчиках. Ср. неудачный (под астериском) перевод: *reading-based assessment tasks* — *«задания для оценки на основе чтения» (Yandex Translate: «оценки чего?»), *helpful to make suggestions for reading-based assessment tasks* — «полезно вносить *предложения по задачам оценки, основанным на чтении» (Yandex Translate: «предложения по задачам оценки чего?»), — который полностью затемняет содержание текста. При этом важно, что многокомпонентные лингвоспецифичные и структурно усложненные термины типа *usage-based grammar* используются не только в соответствующих специальных текстах, чаще всего междисциплинарных, но и в аннотациях к ним, а также, что очень показательно, в названиях научных публикаций, выражая их основное содержание. См. примеры (1) — (4) (здесь и далее курсив и все **выделения** примеров мои. — *Н. Р.*):

1) Van Oost, A., Willems, A., De Sutter, G. Asymmetric syntactic patterns in German-Dutch translation: A *corpus-based* study of the interaction between normalisation and shining through. *International Journal of Translation*. 2016. 28(1). 7–25.

2) Dunn, J. Exposure and emergence in *usage-based grammar*: computational experiments in 35 languages *Cognitive Linguistics*, vol. 33, no. 4, 2022, 659–699. <https://doi.org/10.1515/cog-2021-0106>.

3) Stoet, Gijsbert. PsyToolkit: A novel *web-based* method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*. 2017. № 44 (1). 24–31.

4) Barking, M., Backus, A., Mos, M. Individual corpus data predict variation in judgments: testing the *usage-based* nature of mental representations in a language transfer setting *Cognitive Linguistics*, vol. 33, no. 3, 2022, 481–519. <https://doi.org/10.1515/cog-2021-0105>.

Так, в названии и в самой статье (4) использованы, в частности, следующие термины, легко включающие в свой состав дополнительные терминологические компоненты, но отсутствующие в словаре или переведенные (и переводимые) «компьютерными переводчиками» неправильно: *needs-adapted* data presentation (using principles for making dictionaries that provide users with limited amounts of structured data from which useful information can be retrieved); *a commercially driven language tool* (that uses a language module based on statistics and is in the process of using Artificial Intelligence (AI) technologies) (пример из [Fuertes-Olivera et al., 2020, с. 91]). В статьях (1)–(3) (выше) также встречаются сходные модели терминов, которые рассмотрены далее.

При этом весьма сложным и нетривиальным вопросом для современных цифровых «помощников» является передача словосочетаний, для которых пока еще не выработан последовательный и «дружественный» алгоритм описания и перевода. Часто неудовлетворительная работа цифрового лексикографического и терминологического ресурса объясняется низким качеством лексикографического описания словосочетаний разных типов и видов, особенно терминологических, и способов обращения к ним: «a low database quality originating from a poor compilation or storage of the lexicographic data that are presented in the pop-up window on demand» [Huang et al., 2021, с. 78]. Так, в (1)–(4) встречаются аналогичные модели, которые: 1) нельзя перевести пословно, 2) нуждаются в выведении специальных формул их перевода, 3) которые должны соответствовать нормам выражения языка перевода. При этом следует иметь в виду самое важное правило, которое нигде не прописано явным образом: типологические различия между русским и английским языком проявляются, в частности, в гораздо более активно развитой системе деривации (и соответствующих значений) в русском языке, нежели в английском, в том числе в системе адъективной деривации (ср. *network-based threats* — «сетевые угрозы», где компонент *-based* подразумевается, имплицитруется, присутствует в неявном виде). См. «цифровые многокомпонентные» термины на английском языке из публикаций (1)–(4) и их качественный — правильный и точный — перевод на русский язык, который не смогла синтезировать ни одна система машинного (автоматического) перевода и который в неявном (имплицитном) виде содержит «эквивалент» английской лексемы *-based* (и ей подобных):

— *usage-based grammars* — «грамматики **употребления**» (ср. Google Translate: *грамматики **использования**);

— *usage-based linguistics* — «лингвистика **употребления**»;

— *emergence in usage-based grammar* — «**инновации** в грамматике **употребления**»;

— *revealing usage-based variation* — «установление вариаций в **употреблении**»;

— *usage-based assumption* — «**практическое** предположение»;

— *in line with a usage-based approach* — «в соответствии с **практическим** подходом»;

— *(fundamental) assumptions in usage-based approaches to language* — «(исходные) предположения относительно **практического** подхода к языку»;

— *usage-based predictions (do not explain all of the variation)* — «**практические** предположения (не объясняют все разнообразие выражения смысла)»;

— *usage-based nature (of mental representations)* — «**принципы выражения** (ментальных репрезентаций)»;

— *L2-induced changes in the L1 (of Germans)* — «**влияние** второго языка на первый»;

— *(a novel) web-based method (for running online questionnaires)* (Stoet, Gijsbert, 2017) — «(инновационный) **интернет-метод** (проведения интернет-опросов)»;

— *The item-based nature (of children's early syntactic development)* — «ориентация на **образец** (при раннем усвоении синтаксиса языка детьми)».

Приведенные примеры на английском языке и их переводы на русский язык ярко демонстрируют асимметрию англо-русских терминологических и словообразовательных явлений, их инновационный и лингвоспецифичный характер, ведущую роль разрешения многозначности для входящих в исходное выражение компонентов при их переводе, учет сочетаемости в языке перевода и т. д. Таким образом, важной причиной неудачной работы в этом отношении автоматических переводчиков выступает тот факт, что в соответствующих двуязычных ресурсах отсутствует учет контекстуального окружения лексикографической единицы (they are not context aware), ее синтаксической функции, позиции данного элемента в коммуникации, возможности омографии, идентификации ее полисемии, «избирательности по отношению к своим синтаксическим (и смысловым) «партнерам» и др. В более широком контексте — отсутствует семантическая и метатекстовая информация о темпоральной, стилистической и т. д. отнесенности текста и его компонентов [Bothma et al., 2022, с. 72].

Так что, «несмотря на международный характер IT-терминологии, механизмы, способы и средства ее формирования в английском и русском языках совсем не одинаковы, что является отражением различных типологических особенностей этих языков» [Станкевич, 2020, с. 8] и совсем не предполагает ее поэлементный и пословный перевод. Терминологи также отмечают, что большинство современных терминов состоит из нескольких компонентов: «Наиболее продуктивным способом терминообразования в современной IT-терминологии является образование многокомпонентных терминологических сочетаний, основу которых составляют термины-ядра, которые могут быть как базовыми, так и принадлежащими к терминологическим системам других наук [Акулинина, 2003, с. 155].

Соответственно, в компьютерной, информационной и цифровой терминологии выделяют два основных типа терминов: термины-лексемы и термины-фраземы. Термины-фраземы («кластеры», см. [Hyland, 2008; Riabtseva, 2022]) правомерно трактовать как наиболее продуктивные единицы номинации, поскольку они обеспечивают однозначность и систем-

ность терминологии, наиболее адекватно отвечают требованиям номинации, наиболее полно передают ряд дифференциальных признаков целостного усложнённого понятия, отражают в своей структуре родовидовые, партитивные и другие парадигматические отношения и т. д. [Бабалова, 2015, с. 6]. В результате многокомпонентные термины-фраземы (кластеры) оказываются лингвоспецифичными, идиоматичными и не поддающимися пословному переводу, они нуждаются в «переводческих преобразованиях», таких как экспликация, импликация (или, по-другому, «симплификация, элиминация и т. п.) и нормализация.

3. Результаты и обсуждение = Results and Discussion

Результаты сравнительного анализа автоматического перевода с английского языка на русский современных многокомпонентных лингвоспецифичных идиоматичных терминов из области цифровой обработки информации на естественном языке (NLP) показывают, что их автоматический перевод нуждается в значительном совершенствовании, а соответствующие операции нуждаются в систематизации. При этом наиболее эффективными приемами их перевода выступают широко признанные переводческие процедуры передачи смысла с одного языка на другой. Это процедуры экспликации, импликации (симплификации) и нормализации.

3.1. Импликация, экспликация и нормализация как переводческие приемы перевода (новой) терминологии = Explication, implication and normalization in translating new terminology

Корпусные исследования в области сопоставления оригинала и перевода свидетельствуют, что перевод и оригинал, независимо от языка оригинала и языка перевода, отличаются друг от друга на лексическом, грамматическом и дискурсивном уровне [Baker, 1993 и мн. др.]. При этом самыми типичными, распространенными и естественными способами перевода с одного языка на другой выступают такие «межязыковые» приемы перевода, как, в первую очередь, добавление (необходимого по смыслу) элемента (которое увеличивает обычно объем перевода по сравнению с оригиналом), а затем элиминация (плеонастического) элемента и «нормализация выражения» — его приведение в соответствие с (грамматическими и др.) нормами языка перевода (см. [Van Oost et al., 2016, с. 7] с последующей библиографией), которое, правда, не всегда достигается в полном объеме. Соответствующие процедуры — экспликация, импликация и нормализация — получили название «переводческих универсалий», которые активно анализируются в различных переводоведческих исследованиях и используются при переводе текстов во всех областях знаний.

3.1.1. Импликация

Процесс импликации при переводе (с исходного английского языка на русский) многокомпонентных лингвоспецифичных идиоматичных терминов из области цифровых технологий по обработке естественного языка был продемонстрирован выше. О том, что в них компонент *-based* обычно оказывается при переводе плеонастическим, говорят еще и многочисленные примеры типа: *microprocessor-based control facilities* — «управляющий вычислительный **микроконтроллер**», *microprocessor-based hardware and software* — «**микросредства** управляющей вычислительной техники» и т. п. Аналогичным образом используются в английском языке (и переводятся, точнее, опускаются при переводе) и такие компоненты, как *(computer)-assisted*, *(computer)-aided*, *(corpus)-oriented*, *(robot)-mediated (communication)*, *(computer)-mediated (discourse)*, *mobile-assisted language learning*, *AI-driven decision making*, *a test-driven development of the lexicographic product*, *iterative processes with test-driven development*, *lexicon-driven approach*, *a lexicography-assisted writing*, *computer-supported collaborative writing*, *goal-oriented design*, *meaning-oriented classification of collocations*, *process-oriented writing support with the digital Writing Aid* и т. п.

3.2. Добавление слова по смыслу: экспликация

Примеры экспликации гораздо более многочисленны и наблюдаются во всех без исключения терминологических областях — во-первых, особо характерны для перевода с английского языка на русский (ввиду общеизвестной компактности и «имплицитности» английского, особенно «научного» языка, см., например, [Halliday, 2006]) — во-вторых, и проявляются в добавлении необходимого по смыслу полнозначного слова, исключающего пословный перевод с одного языка на другой — в-третьих:

- *batch processing* — «пакетный **режим** обработки **данных**»;
- *working software* — «рабочее **программное обеспечение**»;
- *refuelling system* — «система **перегрузки топлива**»;
- *engineered safety features* — «**инженерно-технические** средства безопасности»;
- *oil storage tank* — «станция **приема и перекачки** топлива»;
- *estimating of work performance* — «**составление смет** на производство работ»;
- *hard disk drive* — «накопитель на **жестких магнитных дисках**»;
- *out-of-core instrumentation* — «**вне реакторные контрольно-измерительные приборы**» и т. д.

3.3. Нормализация

Нормализация представляет собой прием перевода, особенно с английского языка на русский, в котором используются как вспомога-

тельные средства языка перевода: предлоги, союзы, порядок слов, изменение числа, падежа, части речи и др., так и собственно лексические: уточнение лексического оформления мысли, лексикализация вспомогательных (служебных) слов или, напротив, грамматикализация лексических значений, принятое в данном языке идиоматичное выражение, не допускающее произвольной замены лексических элементов, и т. д. с целью придания естественного звучания (*native-like*) переведенной на другой язык конструкции, ср.:

— *revealing usage-based variation* — «установление вариаций в употреблении»;

— *human error* — «ошибка персонала»;

— *in-core noise diagnostics* — «внутриреакторная шумовая диагностика»;

— *research and development works* — «научно-исследовательские работы»;

— *resistance temperature transducer* — «термопреобразователь сопротивления».

С нормализацией переводного соответствия часто связано использование характерного, «идиоматичного» для языка перевода выражения, ср.: *IT-based phraseology learning* — *изучение фразеологии на основе ИТ (*Google Translate: неточно, дословно*). Ср. пример «нормализации» выражения: *цифровые технологии изучения фразеологии*; а *lexicography-assisted communication* — «коммуникация *с помощью лексикографии» (*Google Translate: неточно*). Ср. пример «нормализации» выражения: *лексикографическая поддержка коммуникации*.

3.4. Автоматический перевод и моделирование естественного интеллекта: преобразование «продуктов» в «процессы» и обратно

Важнейший и самый надежный способ корректного перевода многокомпонентных лингвоспецифичных и потому идиоматичных терминологических конструкций на другой язык — это их включение в автоматические лексикографические и терминологические ресурсы целиком. Так, в исследовании [Fuertes-Olivera, 2022] на основе этих принципов, а также на основе большой предварительной работы излагаются перспективы создания автоматического словаря, совместимого с технологиями ИИ — с автоматической обработкой естественного языка (Natural Language Processing (NLP), которая носит динамический и модульный характер, отличается параллелизмом, динамичностью и высокой коннективностью, свойственной сознанию человека. В качестве «продуктов» коммуникации он содержит целостные единицы (*holistic entities*), которые являются как словами, так и словосочетаниями, ср. [Bergenholtz et al., 2014]. Они кодируются и вклю-

чаются в словари — с одной стороны, и моделируют «процессы», которые отражают их постоянную эволюцию и адаптацию к текущей коммуникативной ситуации, — с другой. Учет соответствующих свойств в цифровых лексикографических ресурсах при их использовании может значительно облегчить интеллектуальную деятельность человека при восприятии и порождении речи с помощью цифровых устройств, а также работу систем ИИ, если создавать лексикографические продукты, которые моделируют такие свойства естественного интеллекта, как преобразование «продуктов» в «процессы» и обратно.

Самым эффективным способом представления многокомпонентных терминов в словаре является их включение в него в полном составе. Это не только облегчает перевод, но и снимает большое количество разнообразных «сопутствующих» проблем, главная из которых — многозначность лексики и «активная» многозначность терминологии. Так, если одиночные лексемы, особенно терминологические, могут иметь от 6 до 25-ти словарных значений, то их использование в составе устойчивых выражений (словосочетаний, кластеров и т. п.) в речи снижает это количество до 1,25 % [Fuertes-Olivera, 2022, с. 126]. Следовательно, чем более представлен на данном ресурсе контекст употребления данного слова (и особенно термина) и точнее охарактеризованы его связи со смежными понятиями, тем меньшее воздействие оказывают на перевод неоднозначность и многозначность слова, которые разрешаются контекстом, ср. *extended units of meaning* [Rundell, 2018]. И чем более подробно описано количество контекстуально связанных употреблений лексических единиц и их производных в данной базе знаний, на данном лексикографическом ресурсе и т. д., тем эффективнее она разрешает их лексическую и грамматическую неоднозначность.

При этом учет производных заглавного слова способствует явной амплификации цифровых лексикографических ресурсов, особенно связанных с процессами NLP. Представление о словаре как о законченном продукте лексикографии трансформируется в представление о ресурсе, который постоянно обновляется, уточняется и «калибруется» — совершенствуется в качестве базы знаний. Важнейшим источником эмпирического материала при этом выступает Интернет [Fuertes-Olivera, 2022, с. 127] как постоянно обновляемая база данных и обязательный источник новых сведений обо всех лексикографических и терминологических единицах.

4. Заключение = Conclusions

Результаты проведенного исследования показывают следующее.

1. Основные свойства современной англоязычной терминологии, особенно в области цифровых технологий обработки информации на есте-

ственном языке, такие как компактность, динамичность, точность, многокомпонентность, межъязыковая асимметрия, идиоматичность, междисциплинарность и др., требуют специальных процедур формализации и автоматизации подбора их межъязыковых соответствий, в том числе при их переводе на русский язык.

2. Самым простым, эффективным и качественным приемом установления соответствующих межъязыковых соответствий выступает их перевод профессиональными переводчиками и последующее их моделирование, формализация, алгоритмизация и автоматизация перевода, а также их включение в полном составе в двуязычные автоматические словари. При этом следует учитывать их способность к расширению, варьированию, модификации, обновлению и т. д.

3. Как и все информационные ресурсы, лексикографические и терминологические базы данных, банки знаний и лексикографические, терминологические и энциклопедические ресурсы должны подвергаться регулярному обновлению, дополнению и совершенствованию в части описания особенностей, сочетаемости, значения и семантического «окружения» терминов — их синонимов, антонимов, аналогов, переводов на другой язык, особенно на основе обновляемых корпусов специальных интернет-текстов и т. д.

4. Таким образом, современные лексикографы и терминографы все более утверждают в мысли, что благодаря беспрецедентному развитию цифровых технологий лексикография и терминография все более сближаются и разрабатывают общие принципы автоматической обработки естественного языка. Так, общие принципы разработки лексических и терминологических дефиниций: их организации, выявления особенностей и отличительных и общих черт — значительно облегчают и упорядочивают процесс создания лексикографических ресурсов, а также их использования в различных образовательных, исследовательских, коммерческих и прочих цифровых проектах, нацеленных на помощь в создании, понимании, переводе или редактировании текстов на естественном языке, а также в их совмещении с разнообразными ИИ-технологиями. В результате оказывается, что точная типология дефиниций лексических и терминологических единиц и тем самым конвергенция терминографии и лексикографии значительно повышают качество работы электронного ресурса, как одноязычного, так и двуязычного. Установлено, что создание строгих дефиниций возможно и даже необходимо не только для терминологической лексики, как считалось ранее, но и для общеупотребительных лексических единиц, что значительно повышает эффективность работы соответствующего двуязычного ресурса [Fuentes-Olivera et al., 2020, с. 90—110; 257—286].

5. Современный этап развития лингвистики отличается от предыдущего переходом от изучения фонологии, синтаксиса и логической семантики к более тщательному анализу функционирования слов в тексте и в речи: к тому, как они используются. В результате лексикография переместилась в центр лингвистических исследований. В то же время сама лексикография преобразуется из традиционного лингвистического описания в новую междисциплинарную науку, все более активно используя корпусные методы исследования и предоставляя широкому кругу специалистов теоретических и прикладных направлений все более точные лексикографические, терминологические, комбинаторные, межъязыковые, методические, стилистические, методологические, переводческие, культурологические, литературоведческие и т. п. сведения об использовании языка в речи и коммуникации, ср. [Hanks et al., 2016, с. 3]. Особо важную роль в развитии современной лексикографии и терминографии играют цифровые технологии, преобразующие их в принципиально новое направление — «цифровая лексикография и терминография», которая ставит перед собой новые задачи и ищет для них новые решения.

Автор заявляет об отсутствии конфликта интересов.

The author declares no conflicts of interests.

Литература

1. Акулинина Т.В. Экстралингвистическая обусловленность особенностей английской терминологии компьютерной информатики [Текст] : диссертация ... кандидата филологических наук : 10.02.04 / Т. В. Акулинина. — Омск, 2003. — 160 с.
2. Бабалова Г. Функциональная природа компьютерной терминологии / Г. Бабалова // Инновационная экономика и общество. — 2015. — № 3 (9). — С. 2—7.
3. Каменева Н. А. Анализ лексических особенностей английского и русского языков в сфере информационных технологий / Н. А. Каменева // Russian Journal of Linguistics. — 2019. — Т. 23. — № 1. — С. 185—199. — DOI: 10.22363/2312-9182-2019-23-1-185-199.
4. Мифтахова Р. Г. Машинный перевод. Нейроперевод / Р. Г. Мифтахова, Е. А. Морозкина // Вестник Башкирского университета. — 2019. — Т. 24. — № 2. — С. 497—502.
5. Пегов С.В. Терминологическая система атомной энергетики (на материале английского языка) : диссертация ... кандидата филологических наук : 10.02.19 / С. В. Пегов. — Тверь, 2018. — 18 с.
6. Пегов С.В. Словарь терминов атомной энергетики : диссертация ... кандидата филологических наук : 10.02.19 / С. В. Пегов. — Курск, 2017. — 163 с.
7. Станкевич Я. А. Особенности перевода современной ИТ-терминологии с английского языка на русский : сопоставительный анализ / Я. А. Станкевич. — Москва : ВШЭ, 2020. — 71 с.
8. Bergenholtz H. The Ideal Number of Lemmas in an Ideal Accounting Dictionary / H. Bergenholtz, B. Norddahl // Hermes. Journal of Language and Communication in Business. — 2014. — № 53. — Pp. 143—150. — DOI:10.7146/hjlc.v27i53.20988.

9. *Bothma Th. J. D.* Information Needs and Contextualization in the Consultation Process of Dictionaries that are Linked to e-Texts / Th. J. D. Bothma, R. H. Gouws // *Lexikos*. — 2022. — № 32 (2). — Pp. 53—81.
10. *Fuertes-Olivera P. A.* The Mental Lexicon in Lexicography : The Dictionarios Valladolid-Uva / P. A. Fuertes-Olivera // *Lexikos*. — 2022. — № 32. — Pp. 118—140. — DOI: <https://doi.org/10.5788/32-1-1712>.
11. *Fuertes-Olivera P. A.* Integrating Terminological Resources in Dictionary Portals : The Case of the <i>Dictionarios Valladolid-UVa</i> / P. A. Fuertes-Olivera, M. Esandi-Baztan // *Lexikos*. — 2020. — № 30 (1). — Pp. 90—110. — DOI: <https://doi.org/10.5788/30-1-1598>.
12. *Fuertes-Olivera P. A.* A Window to the Future : Proposal for a Lexicographically-assisted Writing Assistant / P. A. Fuertes-Olivera, S. Tarp // *Lexicographica*. — 2020. — № 36. — Pp. 257—286.
13. *Hanks P.* International Handbook of Modern Lexis and Lexicography / P. Hanks, G-M. de Schryver (Eds). — Springer Berlin, 2016. — 900 p.
14. *Halliday M. A. K.* Language of Science / M. A. K. Halliday. — London : Bloomsbury Academic, 2006. — 243 c.
15. *Hyland K.* Academic clusters : Text patterning in published and postgraduate writing / K. Hyland // *International Journal of Applied Linguistics*. — 2008. — № 18 (1). — Pp. 41—62. — DOI: <https://doi.org/10.1111/j.1473-4192.2008.00178.x>.
16. *Huang F.* Dictionaries Integrated into English Learning Apps : Critical Comments and Suggestions for Improvement / F. Huang, S. Tarp // *Lexikos*. — 2021. — № 31. — Pp. 68—92.
17. *Kilgarrieff A.* Word from the Chair / A. Kilgarrieff // *Dictionary Writing Systems (DWS). Proceedings of the Fourth International Workshop. 5 September 2006*. — Turin : PreEURALEX, 2006. — 57 p.
18. *Nielsen S.* The Future of Dictionaries, Dictionaries of the Future / S. Nielsen // *The Bloomsbury Companion to Lexicography*. — 2013. — Pp. 355—372. — DOI: [10.5040/9781472541871.ch-018](https://doi.org/10.5040/9781472541871.ch-018).
19. *Riabtseva N. K.* Contemporary Terminological Problems in a Cross-linguistic Perspective / N. K. Riabtseva // *Nauchnyi dialog*. — 2022. — № 11 (6). — Pp. 123—139. — DOI: [10.24224/2227-1295-2022-11-6-123-139](https://doi.org/10.24224/2227-1295-2022-11-6-123-139).
20. *Rundell M.* Searching for Extended Units of Meaning — and What To Do When You Find Them / M. Rundell // *Lexicography*. — 2018. — № 5. — Pp. 5—21. — DOI: <https://doi.org/10.1007/s40607-018-0042-1>.
21. *Tarp S.* Turning Bilingual Lexicography Upside Down : Improving Quality and Productivity with New Methods and Technology / S. Tarp // *Lexikos*. — 2022. — № 32. — Pp. 66—87.
22. *Van Oost A.* Asymmetric syntactic patterns in German-Dutch translation : A corpus-based study of the interaction between normalisation and shining through / A. Van Oost, A. Willems, G. De Sutter // *International Journal of Translation*. — 2016. — № 28 (1). — Pp. 7—25.

Статья поступила в редакцию 20.08.2023,
одобрена после рецензирования 09.10.2023,
подготовлена к публикации 20.10.2023.

References

- Akulina, T. V. (2003). *Extralinguistic conditionality of the features of the English terminology of computer informatics [Text]*. PhD Diss. Omsk. 160 p. (In Russ.).
- Babalova, G. (2015). Functional nature of computer terminology. *Innovative economy and society*, 3 (9): 2—7. (In Russ.).
- Bergenholtz, H., Norddahl, B. (2014). The Ideal Number of Lemmas in an Ideal Accounting Dictionary. *Hermes. Journal of Language and Communication in Business*, 53: 143—150. DOI:10.7146/hjle.v27i53.20988.
- Bothma, Th. J. D., Gouws, R. H. (2022). Information Needs and Contextualization in the Consultation Process of Dictionaries that are Linked to e-Texts. *Lexikos*, 32 (2): 53—81.
- Fuertes-Olivera, P. A., Tarp, S. (2020). A Window to the Future: Proposal for a Lexicographically-assisted Writing Assistant. *Lexicographica*, 36: 257—286.
- Fuertes-Olivera, P. A., Esandi-Baztan, M. (2020). Integrating Terminological Resources in Dictionary Portals: The Case of the <i>Diccionarios Valladolid-UVa</i>. *Lexikos*, 30 (1): 90—110. DOI: <https://doi.org/10.5788/30-1-1598>.
- Fuertes-Olivera, P. A. (2022). The Mental Lexicon in Lexicography: The Diccionarios Valladolid-Uva. *Lexikos*, 32: 118—140. DOI: <https://doi.org/10.5788/32-1-1712>.
- Halliday, M. A. K. (2006). *Language of Science*. London: Bloomsbury Academic. 243 p.
- Hanks, P., De Schryver, G-M. (2016). *International Handbook of Modern Lexis and Lexicography*. Springer Berlin. 900 p.
- Huang, F., Tarp, S. (2021). Dictionaries Integrated into English Learning Apps: Critical Comments and Suggestions for Improvement. *Lexikos*, 31: 68—92.
- Hyland, K. (2008). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics*, 18 (1): 41—62. DOI: <https://doi.org/10.1111/j.1473-4192.2008.00178.x>.
- Kameneva, N. A. (2019). Analysis of lexical features of English and Russian languages in the field of information technology. *Russian Journal of Linguistics*, 23 (1): 185—199. DOI: 10.22363/2312-9182-2019-23-1-185-199. (In Russ.).
- Kilgariff, A. (2006). Word from the Chair. In: *Dictionary Writing Systems (DWS). Proceedings of the Fourth International Workshop. 5 September 2006*. Turin: PreEURALEX. 57 p.
- Miftakhova, R. G., Morozkina, E. A. (2019). Machine translation. Neurotransmission. *Bulletin of Bashkir University*, 24 (2): 497—502. (In Russ.).
- Nielsen, S. (2013). The Future of Dictionaries, Dictionaries of the Future. *The Bloomsbury Companion to Lexicography*. 355—372. DOI: 10.5040/9781472541871.ch-018.
- Pegov, S. V. (2017). *Dictionary of atomic energy terms*. PhD Diss. Kursk. 163 p. (In Russ.).
- Pegov, S. V. (2018). *Terminological system of atomic energy (based on the material of the English language)*. PhD Diss. Tver. 18 p. (In Russ.).
- Riabtseva, N. K. (2022). Contemporary Terminological Problems in a Cross-linguistic Perspective. *Nauchnyi dialog*, 11 (6): 123—139. DOI: 10.24224/2227-1295-2022-11-6-123-139.
- Rundell, M. (2018). Searching for Extended Units of Meaning — and What To Do When You Find Them. *Lexicography*, 5: 5—21. DOI: <https://doi.org/10.1007/s40607-018-0042-1>.
- Stankevich, Ya. A. (2020). *Features of the translation of modern IT terminology from English into Russian: comparative analysis*. Moscow: HSE. 71 p. (In Russ.).



- Tarp, S. (2022). Turning Bilingual Lexicography Upside Down: Improving Quality and Productivity with New Methods and Technology. *Lexikos*, 32: 66—87.
- Van Oost, A., Willems, A., De Sutter, G. (2016). Asymmetric syntactic patterns in German-Dutch translation: A corpus-based study of the interaction between normalisation and shining through. *International Journal of Translation*, 28 (1): 7—25.

*The article was submitted 20.08.2023;
approved after reviewing 09.10.2023;
accepted for publication 20.10.2023.*