РЯБЦЕВА НАДЕЖДА КОНСТАНТИНОВНА доктор филологических наук заведующая отделом прикладной лингвистики Институт языкознания РАН «Дополненная лингвистическая реальность» и современная терминография: межъязыковой аспект Аннотация: В статье обсуждаются новые фундаментальные задачи, связанные с использованием лингвистических знаний и лексикографических принципов их представления в новых видах цифровых ресурсов, использующих технологии искусственного интеллекта в системах «дополненная лингвистическая реальность» класса Writing Assistants (WA), Augmented Writing (AW) и др. Показано влияние цифровых технологий на создание лексикографических двуязычных, особенно терминологических ресурсов, а также их способность к интеграции с формализованными лингвистическими и энциклопедическими знаниями. Ключевые слова: «дополненная лингвистическая реальность», терминография, цифровые ресурсы, новые цифровые технологии, перевод. Augmented Linguistic Reality and Contemporary Terminography: A Cross-linguistic Perspective Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 383 Abstract: The paper exposes new fundamental problems provoked by using linguistic knowledge and its lexicographic representation in advanced bilingual resources based on Artificial Intelligence technologies, such as Writing Assistants (WA), Augmented Writing (AW) et al. The paper tracks the way how bilingual lexicographic translation-oriented terminological resources are created and their «capacity» to integrate with formalized linguistic and encyclopeadic knowledge. Keywords: augmented linguistic reality, terminography, digital resources, new digital technologies, translation. Владимир Григорьевич Гак и современная лексикография. В.Г. Гак был замечательным лингвистом и уникальным лексикографом. Его «Французско-русский словарь активного типа» [1] стал образцом совмещения обширной лингвистической и энциклопедической, в первую очередь, культурологической информации. Особое внимание в «Словаре» уделено также сочетаемости, наиболее лингвоспецифичному аспекту любого языка. Этот словарь прекрасно вписался в «новую лексикографическую культуру» (new lexicographic culture [9, c. 155]) и даже опередил ее, показав, что словарь может давать не только «традиционную» лингвистическую информацию, но и принципиально новую – энциклопедическую. Кроме того, «Словарь» В.Г. Гака обладает еще одним уникальным свойством – «многофункциональностью», которая, как и другие достоинства указанного словаря, особенно ценятся в современной цифровой лексикографии и терминографии [13]. Так, идеи многофункциональности и совмещения лингвистической и экстралингвистической, энциклопедической информации, развитые В.Г. Гаком в области лексикографии, были «подхвачены» современными разработчиками автоматизированных цифровых продуктов: обучающих родному или иностранному языку систем класса (1) Computer-assisted language learning (CALL) [6, c. 3-11; 16], а также систем нового типа – (2) Writing Assistants, Augmented Writing и др. класса «искусственный интеллект» / «дополненная лингвистическая реальность» [17; 3], предназначенных для активной поддержки порождения текста на родном / иностранном языке (text producing). В результате современные лексикографы подчеркивают, что их «продукция» сегодня представляет собой не обычный словарь, а сложно устроенную базу лексикографических данных (БД), которая может быть определенным образом интегрирована в различные цифровые лингвистические инструменты, платформы, порталы и службы («Decisions that have to be taken to prepare a database to feed both this tool and a series of online dictionaries» [14, с. 88]). В этом случае она должна отвечать определенным требованиям, главными из которых выступают многофункциональность, совмещение лингвистической и энциклопедической информации и др. Приоритетные направления современной лексикографии и терминографии. В настоящее время приоритетными направлениями современной лексикографии и терминографии становятся качественные приложения по переводу с родного языка на английский (и обратно), которые Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 384 особо востребованы в Скандинавии, на Пиренеях, в Китае, Юго-восточной Азии и, по сути, во всех других регионах, где английский язык не является родным. Более того, важнейшим стимулом к освоению английского языка, в том числе и при помощи мобильных устройств, становятся беспрецедентные темпы развития самих цифровых технологий, которые не просто внедряются и используются во всех областях научной и практической жизни, но и порождают большое количество специальных терминов [4], которые далеко не сразу попадают в словари и которые появляются первоначально именно в английском языке, представляя собой основные трудности в переводе, в первую очередь, с английского языка на родной, а также с родного на английский. Так, специалисты постоянно подчеркивают, что современная, особенно двуязычная лексикография, особенно в области терминографии и даже в области межъязыковых онлайн-ресурсов, не отвечает возрастающим потребностям ни переводчиков, ни автоматизированных систем перевода (АП) в качественных современных терминологических ресурсах [8, с. 55; 15]: «Постоянная недостаточность терминологии, отстающей от реального технического прогресса, представляет собой архаизирующую сущность любой традиционной терминологии» [5, с. 32]. По целому ряду причин особо остро этот вопрос встает в процессе межъязыкового перевода современной терминологии в области цифровых технологий и обработки лингвистических данных. Современная терминология и ее перевод. При переводе научных публикаций, их названий и аннотаций к ним с русского языка на английский, особую трудность представляет собой современная терминология. Ведь в настоящее время, независимо от того, какой иностранный язык является для нас первым – французский, испанский, немецкий и т. п.., мы обязаны включать в свои публикации перевод их названий, аннотаций и ключевых слов к ним на английский язык. При этом самой большой проблемой выступает то, что в переводах с русского языка на английский почти совсем не встречаются предельно характерные для современного английского научного стиля развернутые лингвоспецифичные термины-словосочетания, возникающие в английском языке лавинообразным образом в результате активного развития цифровых технологий. Они порождают большое количество новых многокомпонентных (уточняющих) терминов во всех без исключения областях знания, включая лингвистику, и в большинстве случаев отсутствуют в автоматических словарях, в автоматических системах перевода и др. цифровых ресурсах и порталах ввиду их инновационности. Их основные типы и модели указаны в [4] и сводятся к следующим (здесь и далее выделены п/ж курсивом): 1) corpus-based (lexicography), languagefocused (research), etc.; 2) subject-specific (terminology), discipline-specific (corpora), context-aware (dictionaries); 3) cross-language (approach); 4) highperformance (tool); 5) (in а) user-friendly (way) и др. В научной коммуникации они обладают целым рядом активных коммуникативных свойств и являются главными носителями предметной информации: Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 385 1. Они способны порождать новые активные лингвоспецифичные терминологические модели, особенно междисциплинарного характера, по аналогии: (1) web-based lexicographic tool, web-supported language learning, a lexicography-assisted writing; meaning-oriented classification of collocations; context-adapted approach; internet-delivered tests; computer assisted pronunciation training; translation-oriented terminography, computer-supported collaborative writing; a suite of internet-delivered diagnostic tests; (2) platform-independent tools, technology-independent content structuring, a space-efficient data structure, computer-adaptive tests of reading, listed in a prioritized, context-aware order, application-consistent checkpoint; (3) a cross-platform ubiquitous language learning service (via mobile phone and interactive television); 4) high-performance lexicographic reference tool; 5) user-friendly interface, in (no) user-unfriendly way. 2. Они способны активно «втягивать» в свой состав окружающий контекст, например: an increasing number of dictionary-based collocation studies; words listed in a prioritized, context-aware order; self-directed L2 vocabulary learning; task-based dictionary use instruction, task-based dictionary training – a detailed demonstration of task-based dictionary training applicable to classroom practice [7, c. 1]; processoriented writing support with the digital Writing Aid и мн. др. 3. Они способны выступать в качестве основного средства организации названия научной публикации, например: «Reference skills or human-centered design» [12]; «Proposal for a lexicography-assisted writing assistant» [11], а также аннотации к ней, и входить в соответствующий список ключевых слов, например: context-aware dictionaries [17, с. 494] и др. Таким образом, стремительное развитие цифровых технологий и соответствующей терминологии и ее исходно «английское» происхождение, а также ее активный коммуникативный и «когнитивный» характер требуют создания лексикографических и терминологических ресурсов нового поколения. Создание лексикографических и терминологических ресурсов нового поколения. Поскольку проблемы перевода терминологии вообще и развернутых терминологических словосочетаний лингвоспецифического характера, возникающих в английском языке в связи с растущими темпами развития цифровых технологий, становятся все более актуальными, то в настоящее время разрабатываются цифровые технологии нового поколения, основанные на принципиально новых лингвистических ресурсах и трендах. Это интегрированные и интегральные словари, лингвистические базы данных, лексикографические порталы, системы класса (3) Artificial Lexicographer, в которых явно выражена тенденция к интеграции лингвистических и энциклопедических данных и которые обладают такими свойствами, как многофункциональность, «контекстуальность», «персонифицированность» и др. [18]: гибкость, настройка на конкретную тематику, на конкретного пользователя, конкретную задачу и т. д. Создаваемая таким образом «дополненная лингвистическая реальность» третьего поколения ориентирована на то, чтобы содержать максимум информации о каждом лингвистическом явлении, и при этом выдавать на экран / по запросу минимальную информацию, наиболее Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 386 соответствующую контексту в широком и узком смысле (alternative, contextadapted approach) [17, c. 496]. При этом появляется возможность обучить учащихся работе со словарями и особенно с многозначными единицами и обеспечить подробное объяснение им связи значения слова с выполняемым тематическим заданием («to provide a detailed demonstration of task-based dictionary training applicable to classroom practice» [7, с. 1]. Кроме того, энциклопедичность, контекстуальность, «персонифицированность» и гибкость особо важны как в терминографии, так и в переводе с одного языка на другой. Так, терминологи подчеркивают, что точные толкования лемм, приближающиеся к терминологическим и энциклопедическим, повышают эффективность работы цифрового устройства в целом и способствуют конвергенции терминологии и лексикографии («Precise definitions, similar to terminological (i.e. encyclopaedic) definitions, for most lemmas increase the tool’s functions» [10, с. 91]). Важнейшей задачей в этом процессе занимает осознание важности разработки удобного дизайна системы и использования доступной пользователю терминологии: the importance of a human-centered design and terminology [14, с. 90]. В целом для современных цифровых лексикографических и терминологических ресурсов нового поколения ставятся задачи создания цифровых порталов, обладающих следующими свойствами: энциклопедичность, многофункциональность, интегральность, «контекстуальность», гибкость, «персонифицированность» и др. Автоматический перевод терминологии. Как указывалось, словари, особенно терминологические, всегда «отставали» от жизни, а в современных «двуязычных» условиях это отставание становится критическим. Более того, межъязыковая асимметрия большинства развернутых терминологических словосочетаний, характерных для развития современных цифровых технологий, приводит к тому, что их автоматический перевод оказывается неприемлемым, главным образом потому, что аутентичные конструкции английского языка в большинстве случаев лингвоспецифичны и носят фразеологизованный характер, как, впрочем, и большинство выражений обыденного языка: «Corpus lexicographers working in the tradition of John Sinclair (of whom the present author is one) argue that electronic dictionaries of the future will have a duty to pay close attention to phraseology and to phraseological meaning in text» [12, с. 109]. В [3; 4] показано, что даже лучшие современные автоматические переводчики типа Google Translate, Yandex Translate и др. не могут правильно перевести даже простейшие трехкомпонентные лингвоспецифичные термины английского языка, например, модели вида commercially driven (tool), corpusbased (approach) и мн. др., на русский язык. Например, естественное для специального текста на английском языке выражение a commercially driven language tool и мн. др. пока еще не включены в терминологические словари, и автоматическими переводчиками переводятся на русский язык буквально, как «коммерческий языковой инструмент», т. е. выражением, которое не используется в специальных текстах на русском языке; ср. правильное: Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 387 «коммерческий лингвистический ресурс / продукт». Еще примеры: АП специальных понятий типа task-based training как «обучение на основе задач», которое по-русски звучит незакончено: «Каких задач?». Это подтверждается использованием данного выражения в целом комплексе терминологических словосочетаний на английском языке и их некорректным АП с английского языка на русский: в результате, например, словосочетание task-based dictionary training по-русски будет иметь 1) слабо терминологический характер: «обучение работы со словарем на основе задач» (Google Translate), 2) требует обязательного объяснения, в чем заключаются задачи. Смысл исходного словосочетания task-based training и подобных ему подразумевает «обучение на компьютере», это «компьютерное обучение», поскольку в их основе лежит исходное понятие – computer-based training – «компьютерное обучение», а словосочетание task-based dictionary training означает «компьютерное обучение работе со словарем». Ср. маркированное «выполнение заданий при обучении работе со словарем». Соответственно, в настоящее время одной из самых актуальных задач в области терминографии выступает создание цифровых интегральных двуязычных ресурсов и лексикографических баз данных, в которых все словарные единицы получают развернутые точные толкования и варианты перевода на иностранный (английский) язык или с английского на родной. При этом самой актуальной задачей выступает перевод на иностранные языки самой современной «цифровой» «развернутой», т.е. состоящей из нескольких компонентов, англоязычной лингвоспецифичной терминологии, которая как никогда быстро развивается и потребности в переводе которой с английского языка постоянно увеличиваются во всех языках. Важно, чтобы соответствующие цифровые ресурсы были хотя бы двуязычными. Ср. «Today new technologies and collaboration between lexicography and information science offer numerous new challenges that can be met by lexicographers» [18, с. 250]. Новые автоматизированные системы перевода строятся на основе указанных выше принципов, призваны учитывать недостатки действующих систем АП и закономерности в переводе терминов специалистом. Они создаются на основе нейронных сетей и принципов самообучения, и призваны преодолеть недостатки действующих систем АП: буквализм, дословность, неидеоматичность и отсутствие принципов формализации и аналогии межъязыковых соответствий. Принципы формализации перевода современных англоязычных многокомпонентных терминологических словосочетаний на русский язык. При формализации установления англо-русских межъязыковых соответствий следует учитывать следующие две противоположные закономерности: 1) для научного стиля на русском языке в принципе характерно дополнение и экспликация понятий, в том числе при переводе с английского языка, поскольку в английском языке научный стиль характеризуется высокой степенью «свернутости» и имплицитности. Так, в «Англо-русском словаре по вычислительной технике и информационным технологиям» [2] (как и во всех других аналогичных специальных словарях) это положение демонстрируют Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 388 соответствия типа: access information – «информация о сеансе связи»; account database – «база данных с информацией о пользователях»; accounting management – «учет работы в сети»; active configuration area – «область памяти для хранения активной конфигурации», batch processing – «пакетный режим обработки данных» (и т.д. и т.п.), в которых выделены явно присутствующие «дополняющие смысл» словосочетания компоненты. Особенности перевода современных англоязычных многокомпонентных терминологических словосочетаний активного типа и лингвоспецифического характера, в первую очередь, связанных с цифровыми технологиями, заключаются в том, что в них часто «нарушается» этот важнейший принцип установления англо-русских соответствий, а именно: «в русском языке устойчивым, компактным и «свернутым» словосочетаниям английского языка чаще всего соответствуют развернутые конструкции с эксплицирующим некоторый встроенный в них смысл компонентом»). Вместо этого, напротив, часть значения термина «морфологизируется» при помощи адъективного аффикса, тем самым упрощая структуру русского термина-эквивалента, ср. network-based threats – сетевые угрозы. 2) В результате при переводе современной многокомпонентной терминологии с английского языка на русский, особенно связанной с компьютерными технологиями, соответствующая ей русская конструкция нередко бывает более компактной благодаря возможности использования адъективных конструкций. Они обладают в русском языке смысловой емкостью и часто не имеют аналога в английском языке – по причине различающихся типологических особенностей этих двух языков, ср. «компьютерное обучение» – computer-based training. В результате рассматриваемые здесь особые лингвоспецифичные аутентичные типы современных развернутых терминологических словосочетаний английского языка переводятся на русский язык более «компактными» адъективными показателями: такими, которые отсутствуют в английском языке – и который выражает в английском языке соответствующие «адъективные» смыслы полнозначными лексемами, ср.: corpus-informed research – «корпусное исследование» corpus-driven analysis – «корпусный анализ» frequency-driven analysis – «частотный анализ» context-adapted approach – «контекстуальный подход» context-adapted analysis– «контекстуальный анализ » commercially driven tool – «коммерческий продукт» discipline-specific corpora – «дисциплинарные корпуса» discipline-specific vocabulary – «дисциплинарный словарь» context-aware dictionaries – «контекстуальные словари» cross-language approach – «межъязыковой подход» high-performance tool – «(высоко)продуктивный инструмент» user-friendly tool – «удобный инструмент» (in a) user-friendly (- unfriendly) way – «удобно (неудобно»). Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 389 Таким образом, в данной статье были использованы современные аутентичные многокомпонентные лингвоспецифичные термины на английском языке и их «расширения» – контексты, которые автоматические переводчики не в состоянии (корректно) перевести на русский язык, и показаны принципы решения соответствующих терминографических и переводческих задач. Сказанное свидетельствует о том, что важнейшей задачей современных цифровых автоматизированных двуязычных лексикографических ресурсов нового поколения, класса «Дополненная лингвистическая реальность», выступает, по началу, включение в их состав, в том числе «целиком», исходных англоязычных многокомпонентных терминов лингвоспецифического характера и их эквивалентов на русском языке, а также учет их способности к вариативности, модификации и расширению. Соответствующая терминология носит «прецизионный» характер и, соответственно, несет основную информационную нагрузку в научной коммуникации и потому требует особо тщательной обработки. Список литературы 1. Гак В.Г. Французско-русский словарь активного типа. Под редакцией В.Г. Гака и Ж. Триомфа. М: Русский язык, 1991. 2. Орлов С.Б. Англо-русский словарь по вычислительной технике и информатике. 4-е изд. М., 2018. URL: https://www.labirint-bookstore.ru/id/92662/ (дата обращения: 01.02.2023) 3. Рябцева Н.К. Лингвистические знания и «дополненная реальность»: перспективы симбиоза // Когнитивные исследования языка / Язык – социальная когниция – коммуникация: материалы XI международного конгресса по когнитивной лингвистике / В.Н. Болдырев (ред.). М.; Тамбов, 2022. № 4 (51). С. 874-878. 4. Рябцева Н. К. Современная лексикография, дополненная лингвистическая реальность и проблемы терминографии // Лингвистика и методика преподавания иностранных языков. М.: ИЯз РАН, 2022. № 1 (16). С. 75–110. 5. Способы номинации в современном русском языке / отв. ред. Д.Н. Шмелев. М.: Наука, 1982. 6. Abel A. Towards a systematic classification framework for dictionaries and CALL // eLexicography in the 21st century: New challenges, new applications. / S. Granger and M. Paquot (eds). Proceedings of eLex 2009. Louvain-la-Neuve, 2009. P. 3-11. 7. Chen Yu. Effect of Learning Conditions on Collocation Gains: A Case Study of Task-based Dictionary Use Instruction // Lexikos. 2022. № 32. P. 1-30. 8. Durán-Muñoz I. Specialised lexicographic resources: a survey of translators’ needs // Cahiers du Cental. 2010. № 7. P. 55-66. Язык и действительность. Научные чтения на кафедре романских языков им. В.Г. Гака. 2023 год 390 9. EURALEX–2022. Dictionaries and Society. Book of Abstracts of the XХII EURALEX International Congress / Klosa-Kückelhaus A. et al. (eds.). Mannheim: IDS-Verlag, 2022. 10. Fuertes-Olivera Pedro A., Esandi-Baztan M.A. Integrating Terminological Resources in Dictionary Portals: The Case of the Diccionarios Valladolid-UVa // Lexikos. 2020. № 30. P. 90-110. 11. Fuertes-Olivera Pedro A., Tarp S. Fourth Industrial Revolution. A window to the future: Proposal for a lexicography-assisted writing assistant // Lexicographica. 2020. P. 257-286. 12. Gouws R., Tarp S. Reference Skills or Human-Centered Design: Towards a New Lexicographic Culture // Lexikos. 2020. № 30. P. 470-498. 12. Hanks P. Elliptical arguments: a problem in relating meaning to use // Cahiers du Cental. 2010. № 7. P. 109-123. 13. Heid U., Gouws R.H. A Model for a Multifunctional Dictionary of Collocations // Atti del XXII Congresso Internazionale di Lessicografia / E. Corino, C. Marello, Cr. Onesti (eds). Torino, Edizioni dell’Orso, 2006 (XII EURALEX International Congress: Proceedings). Vol. 1. P 979-988. 14. Huang F., Tarp S. Dictionaries Integrated into English Learning Apps: Critical Comments and Suggestions for Improvement // Lexikos. 2021. № 6 (31). P. 68-92. 15. Mobile-Assisted Language Learning: A Duolingo Case Study / Loewen S. et al. // ReCALL. 2019. № 31(3). P. 293-311. 16. Pajzs J. On the possibility of creating multifunctional lexicographic databases // Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographic Tools Tomorrow / H. Bergenholtz, S. Nielsen and S. Tarp (eds). Bern: Peter Lang, 2009. Р. 327-354. 17. Tarp S., Fisker K., P. Sepstrup. L2 writing assistants and context-aware dictionaries: New challenges to lexicography // Lexikos. 2017. № 27 (1). P. 494-521. 18. Tarp S., Gouws R. H. Lexicographic Contextualization and Personalization: A New Perspective // Lexikos. 2019. № 29. P. 250-268. 19. Tarp S. Turning Bilingual Lexicography Upside Down: Improving Quality and Productivity with New Methods and Technology // Lexikos, 2022. № 32. P. 66– 87. СТЕПАНЮК ЮЛИЯ ВАЛЕРЬЕВНА кандидат филологических наук, доцент Московский государственный университет имени М.В. Ломоносова Роль фразеологических интернационализмов в обучении фразеологии французского языка Аннотация: В последние годы активно говорят о необходимости формирования фразеологической компетенции у обучающихся иностранном