

<https://www.doi.org/10.33910/1992-6464-2022-204-209-220>

*А. А. Виландеберк*

## ЛИНГВИСТИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ПРОЦЕССА ЦИФРОВИЗАЦИИ ОБРАЗОВАНИЯ: ПРОБЛЕМЫ И ПЕРСПЕКТИВЫ

*Лингвистическое обеспечение процесса цифровизации образования включает комплекс лингвистических и информационных образовательных ресурсов, электронные терминологические базы данных и знаний, корпуса текстов, системы обработки текстов, а также различные сервисы компьютерной поддержки работы всех участников образовательного процесса. Приоритетными для процесса цифровизации образования становятся лингвотехнологии, обеспечивающие, в первую очередь, цифровую репутацию университета, а также структурирование методической составляющей образовательного и исследовательского контента.*

**Ключевые слова:** лингвистическое обеспечение, лингвистические и информационные образовательные ресурсы, лингвотехнологии, цифровизация образования

*A. Vilandeberk*

## LINGWARE FOR THE DIGITALISATION OF EDUCATION: CHALLENGES AND PROSPECTS

*Lingware for the digitalisation of education includes a complex of linguistic and information educational resources, electronic databases of knowledge and terminology, text corpora, text processing systems, as well as various computer support services to facilitate all the stakeholders in education in their activity. Linguistic technologies are a priority in the digitalisation of education. They will ensure, above all, the digital reputation of the university, and effectively manage the methodological resources used in education and research.*

**Keywords:** lingware, linguistic and information educational resources, linguistic technologies, digitalisation of education

### Введение

Задачей статьи является рассмотрение комплекса лингвистических проблем, возникающих при реальной цифровизации образования. Лингвистическое обеспечение процессов трансформации высшего образования (см. [17]) способствует симбиозу образовательного процесса с научно-исследовательской и инновационной деятельностью. С этой точки зрения приоритетным становится вопрос разработки и определения необходимого и достаточного набора лингви-

стических ресурсов и технологий для использования преподавателями, обучающимися и специалистами по формированию и размещению различного рода контента для решения задач по развитию и внедрению современных процессов распространения знаний.

### Лингвистическое, технологическое и методическое обеспечение процессов цифровизации

Двухлетний опыт реализации образовательных программ в дистанционном формате

на филологическом факультете РГПУ им. А. И. Герцена выявил необходимость разработки комплексного подхода не только к определению принципов, методов и технологий дистанционного обучения, но и к проектированию и совершенствованию (структуризации, форматированию, корректровке) всех его составляющих.

Проведенное исследование позволяет выделить три наиболее важных компонента, решение которых дает возможность оптимизировать и повысить эффективность дистанционного обучения: лингвистический, технологический и методический.

Лингвистическое обеспечение процесса цифровизации образования включает комплекс лингвистических и информационных образовательных ресурсов, электронные терминологические базы данных и знаний, корпуса текстов, системы обработки текстов, а также различные сервисы компьютерной поддержки работы всех участников образовательного процесса. Лингвистический компонент для обеспечения продуктивной работы преподавателя, студента, исследователя, специалиста по работе с контентом набора доступных и удобных сетевых сервисов и приложений, информационных и лингвистических образовательных ресурсов и лингвистических технологий для обеспечения профессиональной и научно-исследовательской деятельности в цифровой среде университета.

Под технологическим компонентом, предполагающим, в первую очередь, развитие цифровой инфраструктуры образовательного учреждения, понимается: материально-техническое оснащение образовательного учреждения, а именно, подключение к высокоскоростному (в идеале — беспроводная локальная сеть шестого поколения Wi-Fi 6) интернету с целью обеспечения бесперебойного функционирования цифровой образовательной среды университета, включая нормальную работу большого количества компьютеров и гаджетов, университетских программ и баз данных, образовательных

онлайн платформ, курсов и сервисов, а также обеспечение доступа к большому количеству образовательных ресурсов, электронным библиотекам и различным цифровым коллекциям; цифровые технологии, обеспечивающие проектирование и реализацию образовательного процесса в дистанционном формате, как уже ставшие привычными, так и технологии виртуальной реальности (AR/VR) и Интернета вещей (Internet of Things), единства сети умных устройств и облачной платформы, работающих в режиме реального времени, напрямую связанного с внедрением машинного обучения и искусственного интеллекта, для решения, например, задач тестирования качества освоения знаний.

Методический компонент предполагает разработку, проектирование и необходимую корректировку всех составляющих образовательного процесса в цифровой среде, а также определение принципов представления и структуризации содержательного наполнения контента образовательной программы. При этом нужно исходить из того, что современная образовательная программа должна быть гибкой и адаптивной, оперативно реагировать на возможные изменения условий ее реализации, иметь полноценную ресурсную поддержку и обеспечивать реализацию индивидуальных образовательных маршрутов и развитие профессионального роста обучающихся [6]. Подготовка электронной документации образовательных программ, обусловленная необходимостью их унификации и регламентации представления в цифровой образовательной среде и осуществления автоматического контроля, потребует не только обоснованного выбора соответствующего программного обеспечения (готового и/или адаптированного или специально разработанного), но и подготовку специалистов, обладающих необходимыми навыками работы с такими оболочками, а также разработку системы тренингов для преподавателей по вводу информации в данные системы и адаптации учебной

и научной информации в требуемых форматах представления [4, с. 100].

В рамках проведенного исследования особый интерес представляет лингвистическая составляющая, так как определение концепции лингвистического обеспечения будет способствовать разработке модели сетевого АРМ преподавателя и исследователя.

### **Структура информационных и лингвистических образовательных ресурсов цифровизации**

Сетевые информационные ресурсы, созданные для решения определенных задач, представляют собой структурированные знания, накопленные во всех областях жизнедеятельности человека, в том числе в науке и образовании. Информационные ресурсы легко систематизируются по датологическим и поисковым основаниям и доступны для многократного использования большим количеством пользователей. Сетевые информационные ресурсы можно условно определить как новую форму хранения и представления информации, основной характеристикой которой является ее доступность в любое время и в любом месте. Информационные образовательные ресурсы входят в состав банка национальных информационных ресурсов и включают поисковые системы и каталоги ресурсов, собрания и коллекции различных документов, открытые образовательные и учебные ресурсы, образовательные порталы и сайты, электронные библиотеки, библиотечные системы и архивы, информационные модели и др. (например, информационные системы «Единое окно доступа к информационным ресурсам», «Единая коллекция цифровых образовательных ресурсов», «Мир энциклопедий», Библиографические базы данных ИНИОН РАН).

Как и большинство современных информационных ресурсов, образовательные ресурсы требуют системы управления на основе специальных технологий, основанной на принципах целевой направленности,

структурного моделирования, целевого и коммуникационного соответствия, многоаспектности, должны иметь удобный пользовательский интерфейс и отвечать задачам обеспечения высокого качества образования на различных этапах. Следует учитывать, что ежедневно в мире появляется огромное количество новой информации, размещаемой в сети Интернет, что требует критического подхода к отбору информационных ресурсов, отвечающих требованиям ценности, актуальности, компактности, достоверности и логичности, как корректных источников учебной и научной информации.

Лингвистические ресурсы — это специальным образом организованные базы данных, которые используются для получения и обработки эмпирической информации на естественном языке и ориентированы на обеспечение профессиональной деятельности: справочные ресурсы (полнотекстовые базы данных и знаний, автоматизированные тезаурусы, онтологии и словари), исследовательские (национальные и проблемно-ориентированные корпуса текстов), лингвистический инструментарий для исследователей и переводчиков (терминологические базы данных и глоссарии), системы и лингвистические средства работы с текстом, системы машинного перевода, обучающие тренажеры, системы тестирования, лингвистические атласы, сайты филологических организаций, а также образовательные и справочные ресурсы по русскому и иностранным языкам. В отличие от информационных, лингвистические ресурсы имеют «принципиальный фактологический характер», т. е. они всегда основываются на языковом материале [19]. Лингвистические ресурсы, как правило, являются многоцелевыми и используются как для получения справочной информации, так и для решения конкретных прикладных научно-исследовательских, а также образовательных, в основном лингводидактических задач.

Эффективность использования лингвистических ресурсов в научных и образова-

тельных целях определяется их доступностью, сохранностью, открытостью, возможностью многократного использования, качеством контента и временными затратами на выполнение поставленной задачи, что обуславливает создание и распространение современных сетевых лингвистических ресурсов одной из важнейших задач в области прикладной филологии. Для удобства специалистов в нашей стране создан «Навигатор информационных ресурсов по языкознанию» [10], представляющий собой перечень лингвистических ресурсов, создаваемых учреждениями РАН, и содержащий сведения не только о традиционных ресурсах, но и электронных, таких как электронные коллекции и библиотеки, информационные системы, справочники и энциклопедии, лингвистические и персональные ресурсы, сайты, социальные сети и медиа, информация о мероприятиях и диссертационных советах и др.

Традиционно лингвистические ресурсы определяются как важнейший инструмент деятельности филолога: ученого, исследователя, преподавателя, студента, переводчика и, в последнее время, лингвотехнолога (*language worker*), т. е. специалиста, владеющего навыками применения лингвотехнологий в образовательном пространстве, научно-исследовательской и научно-технической сферах [5, с. 36]. Однако процессы глобализации образования и расширения открытого образовательного пространства обуславливают необходимость владения навыками работы с лингвистическими ресурсами и преподавателями высшей школы. В условиях цифровизации образования преподаватель вынужден наращивать уровень владения навыками медийно-информационной грамотности, являющихся неотъемлемой частью т. н. «навыков XXI века» (*soft skills*), с целью формирования персональной образовательной среды и ресурсной базы методического обеспечения образовательного процесса.

Опыт реализации дистанционного формата обучения выявил целый спектр задач, решение которых позволит не только его

оптимизировать, но и предложить новые формы и способы организации с учетом предъявляемых на сегодняшний день требований, а именно скорости, практичности и индивидуальной направленности. Одной из приоритетных является задача создания специализированного лингвистического обеспечения и сопровождения всего применяемого комплекса методов и технологий.

Лингвистическое обеспечение конкретного процесса или системы предполагает специально сформированный набор лингвистических ресурсов и технологий, определяющих правила, принципы и способы представления, использования и формализации информации в цифровой среде, сформированный для предметной области (в данном случае — образования) для осуществления коммуникации «человек — компьютер — человек». Необходимость установления комплекса лингвистических технологий обусловлена потребностью унификации принципов структурирования информации в образовательной среде учебного заведения для эффективного взаимодействия всех субъектов образовательного и научно-исследовательского процессов, включая онлайн-платформы и базы данных, цифровые образовательные ресурсы, системы искусственного интеллекта и машинного перевода.

Лингвистические технологии, являющиеся частью концепции лингвистического автомата, выдвинутой профессором Р. Г. Плотниковым [14], — это информационные технологии работы с текстом в области естественного языка, обеспечивающие автоматическую переработку текста, функциональным назначением которых становится его адаптация и форматирование в различных языках и системах, стилях и сферах коммуникации (устной, письменной, электронной), организация его структуры, представления, использования и стандартизации.

При установлении состава и структуры лингвистического обеспечения образовательного процесса необходимо определить, какие именно лингвистические технологии могут

и должны использоваться для решения как собственно лингвистических задач в цифровом формате (создание и структурирование текста, компрессия, извлечение и представление специальной информации и знаний, перевод), так и задач разработки специализированных сетевых образовательных ресурсов, систем автоматической переработки текста, программного обеспечения образовательного и научно-исследовательского процессов [3, с. 68].

При этом следует понимать, что с развитием процесса цифровой трансформации образования и цифрового развития общества, спектр используемых для лингвистического обеспечения технологий несомненно будет расширяться.

В прикладной лингвистике принято выделять лингвистические технологии обеспечения поиска и извлечения информации, машинного перевода, генерации текста, автоматической обработки текста, автоматического распознавания речи, автоматического извлечения данных из текста, автоматического реферирования текстов, создания и использования национальных и предметно-ориентированных корпусов текстов (технологии корпусной лингвистики), разработки лексикографических ресурсов [1]. В современном обществе лингвотехнологии становятся универсальными для использования специалистами в любой предметной области знаний.

### **Машинный перевод при создании лингвистических ресурсов**

Формирование и развитие глобального информационного общества способствует распространению и обмену информацией между пользователями на всех естественных языках, однако следует констатировать, что большинство материалов размещается в сети на английском языке, который признан сегодня не только языком международного общения, но и основным языком международного информационного пространства: 70–80% научных работ публикуются

на английском языке, английский является официальным языком международных организаций, а также самым изучаемым в образовательных учреждениях различных типов и уровней [8]. В области образования вся основная информация об образовательном учреждении, образовательных программах, условиях обучения и т. д. публикуется на официальных сайтах вузов на родном и английском языках, что, безусловно, способствует формированию и развитию Интернет-бренда или цифрового брендинга университета. В условиях межъязыковой коммуникации особенно актуальным для преподавателей, студентов и специалистов, осуществляющих содержательное наполнение контента образовательной среды, становится как достаточно свободное владение английским языком в сфере профессиональной коммуникации, так и умение пользоваться технологиями машинного перевода для оперативного извлечения знаний и решения конкретных образовательных и исследовательских задач. Компьютерные системы машинного перевода (МТ) предназначены, в основном, профессиональным переводчикам, владеющим исходным и выходным языками [22; 25; 26; 27; 28].

Следует признать, что большинство компаний, занимающихся разработкой систем машинного перевода, в настоящее время ориентировано как на коммерческого пользователя (бизнес-версии), так и на обычного потребителя (масс-маркет). Безусловно, самыми простыми, известными и доступными для пользователя (преподавателя, студента, ученого) являются автоматические переводчики информационных гигантов — Google (с 2006 года) и Yandex (с 2011 года) для русскоязычного сегмента, использующие технологию нейронного машинного перевода [23; 29; 30] и предназначенные, в первую очередь, для передачи основного смысла оригинального текста. Эти программы масс-маркета адаптивны для всех операционных систем (Windows, MacOS, Linux) и реализованы в формате мобильных приложений на базе Android и iOS.

К достоинствам онлайн-переводчиков Google и Yandex можно отнести скорость выполнения задачи, удобство пользовательского интерфейса, бесплатный доступ, а также практически неограниченные коммерческий и интеллектуальный потенциалы для их модернизации и маркетингового продвижения. Оба переводчика предлагают примерно одинаковое качество перевода (Яндекс — 97, Google — 108 языков), имеют практически одинаковый функционал (перевод текстов из документов, голосовой ввод текста, озвучивание оригинального и переведенного текстов, демонстрацию примеров приведенных слов и др.), используют английский в качестве метаязыка и осуществляют перевод на основе статистического анализа. Недостатки данных онлайн-переводчиков также идентичны: небольшие фрагменты для перевода (Yandex — 10000, Google — 5000 знаков), необходимость серьезной редактуры переводного текста, наличие грамматических ошибок, отсутствие контента, что зачастую ведет к полной или частичной потере смысла, а также возможность добавления варианта перевода слова или словосочетания другим пользователем (Google).

Одним из ведущих игроков на рынке систем машинного перевода является российская компания PROMT (с 1991 года), осуществляющая свою деятельность в области лингвистических технологий и автоматизации переводческих процессов. В отличие от онлайн-переводчиков компания предлагает также и оффлайн переводчики для корпоративного сегмента, работа с которыми обеспечивает конфиденциальность информации — ключевой фактор для отдельных отраслей бизнеса и экономики. Бизнес-версии предлагаемых продуктов обеспечены специализированными терминологическими словарями различных профессиональных областей: банки и финансы, металлургия, медицина и фармацевтика, энергетика, госсектор, IT и телекоммуникации, машиностроение, нефть и газ, образование. В 2010 году компания представила технологию перевода

DeepHybrid, совмещающую классические методики машинного перевода и статистические методы; продукт компании PROMT 18 Master предназначен для работы в автономном режиме и позволяет переводить документы различных форматов с сохранением форматирования; версия PROMT Translation Server 19 предназначена для решения множества задач по работе с данными на разных языках (40+ языков) и поддерживает нейронные технологии перевода. 14 сентября 2021 года компания провела презентацию линейки новых продуктов 22 поколения решений для машинного перевода (в том числе облачный интерфейс Neural Translation Server Developer Edition), работающих на основе нейросетевой технологии, что позволяет получить на выходе высокоточный машинный перевод, практически не требующий дополнительного редактирования и максимально приближенный к «человеческому» переводному тексту.

Современные технологии машинного перевода, обеспечивая достаточно точный высококачественный перевод текста с соблюдением терминологии (технология Smart Neural Dictionary компании PROMT) и оригинальной стилистики, способствуют пониманию и созданию письменного специального текста, локализации учебного и научного контента, а также представлению информации в цифровой образовательной среде.

#### **Гармонизация терминологии лингвистического обеспечения**

Функционирование цифровой образовательной среды университета обеспечивается, в том числе, в ходе решения задач управления операциями с терминологией (*Terminology Management*), т. е. действий по выявлению и фиксации терминологии, гарантирующих использование субъектами корректных терминов во всех документах и материалах [2]. Владение терминологией предметной области «высшее образование» становится необходимым условием успешной профессио-

нальной коммуникации преподавателя, так как именно термины представляют и фиксируют знания предметной области. К сожалению, «среднестатистический» вузовский преподаватель, увлеченный преподаванием и прекрасно владеющий терминологией преподаваемой предметной области знаний, зачастую очень слабо владеет базовой терминологией высшего образования, что приводит к непониманию, а иногда и ошибкам (например, остается неясной разница между значениями терминов *«текущая и промежуточная аттестация»*, *«зачетно-экзаменационная и аттестационная ведомость»*, *«программа дисциплины и рабочая программа дисциплины»* и др.). Кроме того, несмотря на то, что подготовка бакалавров и магистров осуществляется в нашей стране с 1994 года, все еще многие преподаватели не совсем отчетливо представляют уровень квалификации бакалавра и магистра.

В условиях цифровизации (помимо четкого определения понятий *«цифровизация»*, *«цифровая трансформация»*) возникает необходимость создания сетевого, доступного, открытого для пополнения глоссария терминологии высшего образования, являющегося одной из составляющих лингвистического обеспечения. Терминология в области высшего образования требует идентификации, гармонизации и, что особенно важно, унификации. Так, например, если в коммуникации «человек — человек» термины *«образовательная программа»*, *«основная образовательная программа»*, *«основная профессиональная образовательная программа»*, *«образовательная программа подготовки»* номинируют одно и то же понятие, и их использование в тексте не вызывает никаких трудностей ни для адресата, ни для адресанта, то автоматические системы, базы данных и системы машинного перевода могут расценивать их как самостоятельные, независимые единицы. Следует отметить, что в результате постоянного реформирования и модернизации системы высшего образования и ставшего уже регулярным приня-

тия новых образовательных стандартов, таких терминов становится все больше и больше («кредит», «зачетная единица», «единица трудоемкости»). Отдельную группу в терминологии образования представляют термины, которые стали массово тиражироваться без уточнения и конкретизации их значения («модуль», «компетенция», «компетентность», «инновация», «стандарт» и др.), что ведет не только к их различным трактовкам, но и возникновению терминологической путаницы. Например, термин «модуль», который используется в нескольких значениях: единица учебного плана, объединяющая несколько дисциплин, направленных на формирование заданного набора компетенций; автономная организационно-методическая структура учебной дисциплины; технологический узел, ориентированный на обучающихся с различным уровнем мотивации и подготовки; искусственная образовательная среда, направленная на оптимизацию и комплексность получения знаний. Таким образом, ситуация, сложившаяся на сегодняшний день с терминологией образования, обуславливает постановку и необходимость решения проблемы четкого определения терминополья, распознавания конкретного значения термина, а также извлечения новых терминов из иноязычных текстов, что обеспечит корректную работу систем автоматической обработки текста и машинного перевода, программного обеспечения и корректного методического сопровождения образовательного процесса.

#### **Технологии корпусной лингвистики в структуре лингвистического обеспечения**

Решение задач терминологии возможно с применением технологий корпусной лингвистики, которые позволяют формировать лексикографические базы данных, а также осуществлять построение и обогащение онтологий, снятия лексической многозначности, информационного поиска, концептуального анализа, автоматического реферирования, тем самым обеспечивая выявление кандида-

тов в термины для включения в терминологические словари, тезаурусы и онтологии [7]. Национальные и специализированные (параллельные, тематические, исследовательские и учебные) корпуса текстов, под которыми принято понимать объединенные по конкретному типологическому признаку, структурированные, размеченные и унифицированные большие массивы текстов, и корпусные менеджеры (программы Ant-Conc, Sketch Engine и др.) востребованы в научной и образовательной деятельности для решения конкретных профессиональных задач в области лингвистики, перевода, терминографии, в социолингвистике и лингвидидактике.

Современные методы корпусной лингвистики, такие как автоматизированное извлечение информации, текстовые поиски (конкордансы и контексты), сравнительный и статистический анализ, позволяют получать точные эмпирические данные при проведении различного рода не только лингвистических, но и литературоведческих, социологических, культурологических исследований, которые могут использоваться и в учебном процессе, в том числе для выполняемых обучающимися выпускных квалификационных и курсовых работ, выполнения различных проектов и др. Традиционно лингвистические корпуса текстов применяются при обучении иностранному языку, однако в последнее время они все больше используются при обучении русскому языку и русскому как иностранному, а также в качестве иллюстративных при преподавании других (в основном гуманитарных) дисциплин — «применение корпусов может дать возможность определенного угла зрения на феномены культуры, истории и социальной психологии» [9].

В других предметных областях самостоятельно созданный преподавателем специализированный предметно-ориентированный корпус текстов, включенный в структуру лингвистического обеспечения, может быть представлен как ресурсная база знаний по конкретной дисциплине и многократно

использован для организации самостоятельной и исследовательской работы обучающихся. Использование лингвотехнологий корпусной лингвистики повышает эффективность работы специалиста-лингвотехнолога, целью которой является поддержание корректного функционирования всего комплекса лингвистического обеспечения цифровой образовательной среды университета, в области разработки, настройки и адаптации автоматизированных систем информационного поиска, машинного перевода и распознавания речи.

Таким образом, в условиях цифровизации сетевые технологии корпусной лингвистики становятся актуальным инструментом профессиональной деятельности всех субъектов образовательной среды университета.

#### **Цифровизация образования: общие проблемы и перспективы**

Развитие и становление экономического потенциала общества обусловили необходимость интеграции образовательного и научного пространства, а также изменение роли высшего образования, которое становится первым звеном инновационной цепочки «образование — исследования — венчурные проекты — массовое освоение инноваций» [18, с. 4].

С одной стороны, в российском высшем образовании накоплен многолетний опыт становления и развития научных школ по фундаментальным направлениям науки, обеспечивающих накопление научного и образовательного потенциала. С другой — есть инерция сведения инновационных и, в частности, информационных технологий к элементарному использованию компьютерной техники без понимания того, какие требования инновации предъявляют к образованию в целом и образовательной среде университета в частности. Как показывают исследования (см., например, [21]), сам термин «*цифровизация образования*» часто понимается просто как информатизация процессов обучения [11], в то время, как



основной задачей цифровизации является обеспечение нового качества организации и реализации образовательного процесса с учетом темпов информационного развития общества и изменения структуры и состава необходимых ему квалификаций.

Соответственно, при организации новых процессов в образовании следует ясно понимать, что четвертая техническая революция («Industry 4.0 / Промышленность 4.0»), подразумевающая активное внедрение информационных технологий в промышленность, автоматизацию различного рода бизнес-процессов и стремительное распространение искусственного интеллекта во всех сферах жизни человека, неизбежно приведет не только к росту конкурентоспособности экономики, но и к пересмотру требований и условий реализации профессиональной подготовки специалистов, способных работать в соответствии с новыми приоритетами и задачами [5, с. 35]. «Цифровизация» традиционно предполагала внедрение технологий и соответствующей инфраструктуры для обеспечения цифровой экономики компетентными специалистами [11] и формирования человеческого капитала для создания высокотехнологичных продуктов современной экономики («Industry 4.0 / Промышленность 4.0»).

### Заключение

Рассмотренные проблемы формирования и структурирования лингвистического обеспечения образовательной среды университета в условиях цифровизации образования могут быть решены в рамках реализации комплексного подхода использования современных принципов и методов прикладной и корпусной лингвистики, потенциала информационно-коммуникационных технологий и возможностей сети Интернет. Для обе-

спечения цифровизации образования необходимо создание высокотехнологичной, легко расширяемой и масштабированной, гибкой и адаптивной по своим функциональным возможностям платформы, обеспечивающей реализацию образовательного и научно-исследовательского процессов; с другой стороны, вовлечение в этот процесс большого количества пользователей, что приведет (и, как показывает опыт, уже привело) к качественному рывку в области образовательного контента, изменению подходов к реализации образовательных программ и появлению новых образовательных методик и технологий.

Приведенные в статье данные в целом характеризуют состав и структуру лингвистического обеспечения, демонстрируют необходимость разработки сетевых лингвистических ресурсов для решения научно-исследовательских и образовательных задач в цифровой образовательной среде. Основное внимание следует уделять использованию современных лингвистических технологий (машинного перевода, терминологизации, корпусной лингвистики) как важнейшего инструмента профессиональной деятельности ученого, исследователя, преподавателя, студента, переводчика и лингвотехнолога по содержательному наполнению образовательного и научного контента. Методологической основой исследования являются концепции лингвистического автомата проф. Р. Г. Пиотровского и АРМ филолога проф. Л. Н. Беляевой.

Дальнейшее направление исследования предполагает а) разработку концепции сетевого АРМ преподавателя вуза; б) определение влияния лингвистического обеспечения и лингвистических технологий на эффективность цифрового брендинга университета.

### СПИСОК ЛИТЕРАТУРЫ

1. Антопольский А. Б. Лингвистические ресурсы и технологии в России: состояние и перспективы // Социальные новации и социальные науки. 2021. № 2 (4). С. 114–131. <https://doi.org/10.31249/snsn/2021.02.08>

2. *Беляева Л. Н.* Лингвистические технологии в современном сетевом пространстве: Language work-er в индустрии локализации. СПб.: Книжный дом, 2016. 134 с.
3. *Беляева Л. Н., Блохина А. В., Камшилова О. Н.* Концепция профессиональной подготовки лингвотехнолога в России // Известия Российского государственного педагогического университета им. А. И. Герцена. 2017. № 183. С. 62–72.
4. *Беляева Л. Н., Камшилова О. Н.* Профессиональная подготовка лингвотехнолога в вузе: зарубежный опыт и новые задачи российского образования // Информационное общество: образование, культура и технологии будущего. 2018. № 2. С. 99–111. <https://doi.org/10.17586/2587-8557-2018-2-99-111>
5. *Беляева Л. Н., Камшилова О. Н.* Проблемы и перспективы профессиональной подготовки лингвотехнолога // International Journal of Open Information Technologies. 2018. Т. 6. № 12. С. 35–42.
6. *Виландеберк А. А., Шубина Н. Л.* Проектирование образовательных программ в условиях формирования новой образовательной среды // Научное мнение. 2020. № 3. С. 33–43. <https://doi.org/10.25807/RVN.2224378.2020.3.33.43>
7. *Захаров В. П., Хохлова М. В.* Автоматическое выявление терминологических словосочетаний // Структурная и прикладная лингвистика. № 10. 2014. С. 182–200.
8. *Камшилова О. Н.* Международный язык современного информационного пространства (Инновационные образовательные технологии в описании лингвистического объекта): научно-методические материалы: для студентов высших учебных заведений, обучающихся по направлению «540300 (050300) Филологическое образование». СПб.: Книжный Дом, 2008. 152 с.
9. *Масевич Ф., Захаров В. П.* Методы корпусной лингвистики в исторических исследованиях // Компьютерная лингвистика и вычислительные онтологии. Сборник научных статей. Труды XIX Международной объединенной научной конференции «Интернет и современное общество», Санкт-Петербург, 22–24 июня 2016 г. [Электронный ресурс]. URL: <https://openbooks.itmo.ru/ru/file/4102/4102.pdf> (дата обращения 17.02.2022).
10. Навигатор информационных ресурсов по языкознанию. [Электронный ресурс]. URL: <http://niryaz.inion.ru/> (дата обращения: 17.02.2022).
11. *Некрасов В. Н.* Инновация, информатизация, цифровизация: соотношение и особенности правовой регламентации // Вопросы российского и международного права. 2018. Т. 8. № 11А. С. 137–143.
12. Программа «Цифровая экономика Российской Федерации» Утверждена распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р. [Электронный ресурс]. URL: <http://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf> (дата обращения 17.02.2022).
13. Паспорт федерального проекта «Цифровая образовательная среда»: Приложение к протоколу заседания проектного комитета по национальному проекту «Образование» от 7 декабря 2018 г. № 3. [Электронный ресурс]. URL: <https://edu-frn.spb.ru/files/iiMBxQ4cNH1BCsaWn2WqDgFinWeU3rVYpmO6sd33.pdf> (дата обращения 17.02.2022).
14. *Пиотровский Р. Г.* Лингвистический автомат (в исследовании и непрерывном обучении). СПб.: Изд-во РГПУ им. А. И. Герцена, 1999. 256 с.
15. Приказ Министерства науки и высшего образования Российской Федерации от 14 марта 2020 г. № 397 «Об организации образовательной деятельности в организациях, реализующих образовательные программы высшего образования и соответствующие дополнительные профессиональные программы, в условиях предупреждения распространения новой коронавирусной инфекции на территории Российской Федерации». [Электронный ресурс]. URL: <https://old.minobrnauki.gov.ru/common/upload/library/2020/03/main/397.pdf> (дата обращения 17.02.2022).
16. Проблемы и перспективы цифровой трансформации образования в России и Китае. II Российско-китайская конференция исследователей образования «Цифровая трансформация образования и искусственный интеллект». Москва, Россия, 26–27 сентября 2019 г. / отв. ред. И. В. Дворецкая. М.: Издательский дом Высшей школы экономики, 2019. 155 с.
17. Программа «Приоритет–2030». [Электронный ресурс]. URL: <https://minobrnauki.gov.ru/action/priority2030/> (дата обращения 17.02.2022).
18. Российское образование–2020: модель образования для экономики, основанной на знаниях: к IX международной научной конференции «Модернизация экономики и глобализация», Москва, 1–3 апреля 2008 г. / под ред. Я. И. Кузьмина, И. Д. Фрумина. М.: Издательский дом ГУ ВШЭ, 2008. 39 с.
19. *Рычкова Л. В.* Языковые и информационные ресурсы: проблемы идентификации и классификации // Карповские научные чтения: сборник научных статей. Вып. 9. В 2 ч. Ч. 1. Минск: Белорусский Дом печати, 2015. С. 27–32.

20. Стариченко Б. Е. Цифровизация образования: иллюзии и ожидания // Педагогическое образование в России. 2020. № 3. С. 49–58. <https://doi.org/10.26170/po20-03-05>
21. Babych B., Buckley A., Hughes R., Babych S. Machine translation technology in advanced language teaching and translator training: A corpus-based approach to post-editing MT output. [Online]. Available at: <http://talc10.ils.uw.edu.pl/files/2011/09/BABYCHBogdan1.pdf> (accessed 17.02.2022).
22. Belinkov Y., Durrani N., Dalvi F., Sajjad H., Glass J. On the linguistic representational power of neural machine translation models // *Computational Linguistics*. 2020. Vol 46. No. 1. P. 1–52. [https://doi.org/10.1162/coli\\_a\\_00367](https://doi.org/10.1162/coli_a_00367)
23. De Almeida G., O'Brien S. Analysing post-editing performance: Correlations with years of translation experience // Proceedings of the 14<sup>th</sup> Annual conference of the European Association for Machine Translation, May 27–28, 2010. Saint Raphael: European Association for Machine Translation Publ., 2010. [Online]. Available at: <https://www.aclweb.org/anthology/2010.eamt-1.19> (accessed 17.02.2022).
24. Escribe M. Human evaluation of neural machine translation: The Case of *deep learning* // Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019) / ed. by I. Temnikova, C. Orasan, G. Corpas Pastor, R. Mitkov. Varna: Incoma Ltd. Publ., 2019. P. 36–46.
25. Ortiz-Martínez D. Online learning for statistical machine translation // *Computational Linguistics*. 2016. Vol. 42. No. 1. P. 121–161. [http://dx.doi.org/10.1162/COLI\\_a\\_00244](http://dx.doi.org/10.1162/COLI_a_00244)
26. Resende N., Way A. MTrill project: Machine translation impact on language learning // Proceedings of the 22<sup>nd</sup> Annual Conference of the European Association for Machine Translation / ed. by A. Martins, H. Moniz, S. Fumega, B. Batista. Geneva: European Association for Machine Translation Publ., 2020. P. 497–498.
27. Tsujii J. Natural language processing and computational linguistics // *Computational Linguistics*. 2021. Vol. 47. No. 4. P. 707–727. [https://doi.org/10.1162/coli\\_a\\_00420](https://doi.org/10.1162/coli_a_00420)
28. Vázquez R., Raganato A., Creutz M., Tiedemann J. A systematic study of inner-attention-based sentence representations in multilingual neural machine translation // *Computational Linguistics*. 2020. Vol. 46. No. 2. P. 387–424. [https://doi.org/10.1162/coli\\_a\\_00377](https://doi.org/10.1162/coli_a_00377)

## REFERENCES

1. Antopol'skij A. B. Lingvisticheskie resursy i tekhnologii v Rossii: sostoyanie i perspektivy // Sotsial'nye novatsii i sotsial'nye nauki. 2021. № 2 (4). S. 114–131. <https://doi.org/10.31249/snsn/2021.02.08>
2. Belyaeva L. N. Lingvisticheskie tekhnologii v sovremennom setevom prostranstve: Language worker v industrii lokalizatsii. SPb.: Knizhnyj dom, 2016. 134 s.
3. Belyaeva L. N., Blokhina A. V., Kamshilova O. N. Kontsepsiya professional'noj podgotovki lingvotekhnologa v Rossii // *Izvestiya Rossijskogo gosudarstvennogo pedagogicheskogo universiteta im. A. I. Gertsena*. 2017. № 183. S. 62–72.
4. Belyaeva L. N., Kamshilova O. N. Professional'naya podgotovka lingvotekhnologa v vuze: zarubezhnyj opyt i novye zadachi rossijskogo obrazovaniya // *Informatsionnoe obshchestvo: obrazovanie, kul'tura i tekhnologii budushchego*. 2018. № 2. S. 99–111. <https://doi.org/10.17586/2587-8557-2018-2-99-111>
5. Belyaeva L. N., Kamshilova O. N. Problemy i perspektivy professional'noj podgotovki lingvotekhnologa // *International Journal of Open Information Technologies*. 2018. T. 6. № 12. S. 35–42.
6. Vilanderberk A. A., Shubina N. L. Proektirovanie obrazovatel'nykh programm v usloviyakh formirovaniya novoj obrazovatel'noj sredy // *Nauchnoe mnenie*. 2020. № 3. S. 33–43. <https://doi.org/10.25807/PBH.22224378.2020.3.33.43>
7. Zakharov V. P., Khokhlova M. V. Avtomaticheskoe vyyavlenie terminologicheskikh slovosochetaniy // *Strukturnaya i prikladnaya lingvistika*. № 10. 2014. S. 182–200.
8. Kamshilova O. N. Mezhdunarodnyj yazyk sovremennogo informatsionnogo prostranstva (Innovatsionnye obrazovatel'nye tekhnologii v opisanih lingvisticheskogo ob'ekta): nauchno-metodicheskie materialy: dlya studentov vysshikh uchebnykh zavedenij, obuchayushchikhsya po napravleniyu “540300 (050300) Filologicheskoe obrazovanie”. SPb.: Knizhnyj Dom, 2008. 152 s.
9. Masevich F., Zakharov V. P. Metody korpusnoj lingvistiki v istoricheskikh issledovaniyakh // *Komp'yuternaya lingvistika i vychislitel'nye ontologii. Sbornik nauchnykh statej. Trudy XIX Mezhdunarodnoj ob'edinennoj nauchnoj konferentsii “Internet i sovremennoe obshchestvo”*, Sankt-Peterburg, 22–24 iyunya 2016 g. [Elektronnyj resurs]. URL: <https://openbooks.itmo.ru/ru/file/4102/4102.pdf> (data obrashcheniya 17.02.2022).
10. Navigator informatsionnykh resursov po yazykoznaniyu. [Elektronnyj resurs]. URL: <http://niryaz.inion.ru/> (data obrashcheniya: 17.02.2022).

11. *Nekrasov V. N.* Innovatsiya, informatizatsiya, tsifrovizatsiya: sootnoshenie i osobennosti pravovoj reglamentatsii // *Voprosy rossijskogo i mezhdunarodnogo prava*. 2018. T. 8. № 11A. S. 137–143.
12. Programma “Tsifrovaya ekonomika Rossijskoj Federatsii” Utverzhdena rasporyazheniem Pravitel’sтва Rossijskoj Federatsii ot 28 iyulya 2017 g. № 1632-r. [Elektronnyj resurs]. URL: <http://static.government.ru/media/les/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf> (data obrashcheniya 17.02.2022).
13. Pasport federal’nogo proekta “Tsifrovaya obrazovatel’naya sreda”: Prilozhenie k protokolu zasedaniya proektnogo komiteta po natsional’nomu proektu “Obrazovanie” ot 7 dekabrya 2018 g. № 3. [Elektronnyj resurs]. URL: <https://edu-frn.spb.ru/files/iiMBxQ4cNH1BCsaWn2WqDgFinWeU3rVYpmO6sd33.pdf> (data obrashcheniya 17.02.2022).
14. *Piotrovskij R. G.* Lingvisticheskij avtomat (v issledovanii i nepreryvnom obuchenii). SPb.: Izd-vo RGPU im. A. I. Gertsena, 1999. 256 s.
15. Prikaz Ministerstva nauki i vysshego obrazovaniya Rossijskoj Federatsii ot 14 marta 2020 g. № 397 “Ob organizatsii obrazovatel’noj deyatel’nosti v organizatsiyakh, realizuyushchikh obrazovatel’nye programmy vysshego obrazovaniya i sootvetstvuyushchie dopolnitel’nye professional’nye programmy, v usloviyakh preduprezhdeniya rasprostraneniya novoj koronavirusnoj infektsii na territorii Rossijskoj Federatsii”. [Elektronnyj resurs]. URL: <https://old.minobrnauki.gov.ru/common/upload/library/2020/03/main/397.pdf> (data obrashcheniya 17.02.2022).
16. Problemy i perspektivy tsifrovoj transformatsii obrazovaniya v Rossii i Kitae. II Rossijsko-kitajskaya konferentsiya issledovatelej obrazovaniya “Tsifrovaya transformatsiya obrazovaniya i iskusstvennyj intellekt”. Moskva, Rossiya, 26–27 sentyabrya 2019 g. / otv. red. I. V. Dvoret’skaya. M.: Izdatel’skij dom Vyshej shkoly ekonomiki, 2019. 155 s.
17. Programma “Prioritet–2030”. [Elektronnyj resurs]. URL: <https://minobrnauki.gov.ru/action/priority2030/> (data obrashcheniya 17.02.2022).
18. Rossijskoe obrazovanie–2020: model’ obrazovaniya dlya ekonomiki, osnovannoj na znaniyakh: k IX mezhdunarodnoj nauchnoj konferentsii “Modernizatsiya ekonomiki i globalizatsiya”, Moskva, 1–3 aprelya 2008 g. / pod red. Ya. I. Kuz’minova, I. D. Frumina. M.: Izdatel’skij dom GU VShE, 2008. 39 s.
19. *Rychkova L. V.* Yazykovye i informatsionnye resursy: problemy identifikatsii i klassifikatsii // *Karpovskie nauchnye chteniya: sbornik nauchnykh statej*. Vyp. 9. V 2 ch. Ch. 1. Minsk: Belorusskij Dom pechati, 2015. S. 27–32.
20. *Starichenko B. E.* Tsifrovizatsiya obrazovaniya: illyuzii i ozhidaniya // *Pedagogicheskoe obrazovanie v Rossii*. 2020. № 3. S. 49–58. <https://doi.org/10.26170/po20-03-05>
21. *Babych B., Buckley A., Hughes R., Babych S.* Machine translation technology in advanced language teaching and translator training: A corpus-based approach to post-editing MT output. [Online]. Available at: <http://talcl0.ils.uw.edu.pl/files/2011/09/BABYCHBogdan1.pdf> (accessed 17.02.2022).
22. *Belinkov Y., Durrani N., Dalvi F., Sajjad H., Glass J.* On the linguistic representational power of neural machine translation models // *Computational Linguistics*. 2020. Vol 46. No. 1. P. 1–52. [https://doi.org/10.1162/coli\\_a\\_00367](https://doi.org/10.1162/coli_a_00367)
23. *De Almeida G., O’Brien S.* Analysing post-editing performance: Correlations with years of translation experience // *Proceedings of the 14<sup>th</sup> Annual conference of the European Association for Machine Translation*, May 27–28, 2010. Saint Raphael: European Association for Machine Translation Publ., 2010. [Online]. Available at: <https://www.aclweb.org/anthology/2010.eamt-1.19> (accessed 17.02.2022).
24. *Escribe M.* Human evaluation of neural machine translation: The Case of deep learning // *Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)* / ed. by I. Temnikova, C. Orasan, G. Corpas Pastor, R. Mitkov. Varna: Incoma Ltd. Publ., 2019. P. 36–46.
25. *Ortiz-Martínez D.* Online learning for statistical machine translation // *Computational Linguistics*. 2016. Vol. 42. No. 1. P. 121–161. [http://dx.doi.org/10.1162/coli\\_a\\_00244](http://dx.doi.org/10.1162/coli_a_00244)
26. *Resende N., Way A.* MTrill project: Machine translation impact on language learning // *Proceedings of the 22<sup>nd</sup> Annual Conference of the European Association for Machine Translation* / ed. by A. Martins, H. Moniz, S. Fumega, B. Martins, F. Batista. Geneva: European Association for Machine Translation Publ., 2020. P. 497–498.
27. *Tsujii J.* Natural language processing and computational linguistics // *Computational Linguistics*. 2021. Vol. 47. No. 4. P. 707–727. [https://doi.org/10.1162/coli\\_a\\_00420](https://doi.org/10.1162/coli_a_00420)
28. *Vázquez R., Raganato A., Creutz M., Tiedemann J.* A systematic study of inner-attention-based sentence representations in multilingual neural machine translation // *Computational Linguistics*. 2020. Vol. 46. No. 2. P. 387–424. [https://doi.org/10.1162/coli\\_a\\_00377](https://doi.org/10.1162/coli_a_00377)