УДК 33

Формирование цели внедрения компьютерной лингвистики для целей форсайта в экономическом развитии

Капитанов Андрей Иванович

Ассистент,
Национальный исследовательский университет
«Московский институт электронной техники»,
124498, Российская Федерация, Зеленоград, площадь Шокина, 1;
e-mail: andrey@kapdx.ru

Аннотация

Компьютерная лингвистика является дисциплиной обязательной для изучения на филологических факультетах, причем именно для переводчиков она составляет значительно больший интерес, поскольку они делают возможным осуществление переводов быстрее и с большим качеством, поэтому переводчик должен быть в курсе ИТ лингвистических аспектов не одного языка, как филолог, а сразу двух, что существенно увеличивает нагрузку. Проблемой данной статьи является освещение того, почему компьютерная лингвистика представляет особый интерес для начинающих, профессиональных переводчиков и студентов специальности «Перевод», поскольку это поможет студенту более эффективно структурировать процесс обучения, а специалисту – дополнить промежутки в имеющихся знаниях. На данный момент самыми распространенными способами применения интеллектуальных компьютерных систем для осуществления автоматизированного письменного перевода работа со словарями, глоссариями, памятью переводов и корпусами. Кроме того, есть специализированное программное обеспечение, применяемое для решения конкретных задач, например, аудиовизуального перевода, локализации продуктов программного обеспечения, редактирование текстов.

Для цитирования в научных исследованиях

Капитанов А.И. Формирование цели внедрения компьютерной лингвистики для целей форсайта в экономическом развитии // Экономика: вчера, сегодня, завтра. 2019. Том 9. № 7А. С. 23-31.

Ключевые слова

Лингвистика, экономическое развитие, форсайт, ИТ, внедрение.

Введение

Несмотря на то, что компьютерная лингвистика как лингвистическая дисциплина возникла и сформировалась лишь во второй половине двадцатого столетия, собственно теме гендерной лингвистики было посвящено большое количество публикаций, в том числе советских [Boguraev, 1994; Coscoy, 1997; Huet, 2003; Shao, 2013], европейских [Lecomte, 2001; Schmidt, 2014], зарубежных [Houben, 2009] и российских [Dahl, 2009; Houben, 2009; Ide, 1992; McDonald, 1991; Simon, 1989; Wu, 2018]. Однако ни один из этих источников не рассматривает вопрос роли, которую играет компьютерная лингвистика для переводчиков, ни содержит сведения, полезные для переводчиков, выдаваемые наиболее важными для привлечения в практическую переводческую деятельность. Это обуславливает актуальность данной статьи. Итак, целью данной статьи является:

1) Прояснить роль дисциплины «компьютерная лингвистика» в обучении профессиональных переводчиков;
2) Назвать причины, по которым знания компьютерной лингвистики является необходимым инструментом в работе переводчика как межкультурного коммуникатору;
3) Доказать необходимость изучения переводчиком определенных областей компьютерной лингвистики и практического применения усвоенного в процессе изучения данной дисциплины теоретического материала.

Основная часть

Первые серьезные известные работы в области компьютерной лингвистики стали появляться еще в пятидесятые годы XX веках [Shao, 2013; Coscoy, 1997]. Обусловлены они были развитием технического прогресса и постепенным привлечением компьютеров – которые на тот момент еще называли ЭВМ (электронно-вычислительная машина) – ко всем отраслям человеческой культуры. Вполне понятно, что лингвисты также стали задумываться, как использовать неизведанный тогда потенциал компьютеров для собственных целей. Мощности ЭВМ для обработки информации в цифровом виде вызвали недоумение уже тогда, поэтому было создано специальное подразделение лингвистики – компьютерная лингвистика, одной из сфер исследования которой был и машинный или автоматизированный перевод. Именно тогда, еще в далекие 50-е годы, появились зародыши эры, которая в будущем оставит переводчиков без работы и сведет их роль в лучшем случае к операторам машины, которая будет использовать алгоритмы для перевода лучше, чем человек смог бы выполнить. Но для того, чтобы лучше понять полноту значения компьютерной лингвистики, следует прояснить сущность этой дисциплины, механизмы включения ее появления и перспективы развития, а также значение ее как для филологов, так и собственно для переводчиков.

Итак, технологическим открытием, которое знаменовало прорыв в виде и принципах функционирования ЭВМ, стало изобретение транзисторов на замену устаревшим электронным лампам как компонент схем компьютеров. Транзистор – это радиоэлектронный компонент, сделанный из полупроводникового материала, который используется для усиления, генерации, коммутации и преобразования электрических сигналов. Сейчас именно транзисторы лежат в основе электронных устройств, микросхем и большинства техники, что содержит электронные компоненты. За свое изобретение, представленное 23 декабря 1947 года, У. Шокли, Д. Бардин и У. Браттейн получили в 1956 году Нобелевскую премию. Транзисторы не только повысили скорость обмена электронными импульсами с компьютером, но и минитюаризировали их,

сделав значительно компактнее, как мы знаем их сейчас. Транзисторы потерпели второй этап эволюции в 1990-е годы, когда конструкция была усовершенствована до «биполярного транзистора», меньше по размеру и значительно мощнее. Сегодня ведутся активные разработки по внедрению полупроводникового материала для транзисторов, что мог бы еще ускорить их работу, предлагают использовать принципиально новый синтетический материал графен. Наука и прогресс в области физической стороны компьютера не стоит на месте, обещая нам новые открытия и возможности.

После изобретения и удивительно быстрого появления компьютеров нового поколения следующим серьезным шагом было изобретение второго компонента схемы, не-физического. Это интеллектуальная схема, которая, собственно, и стоит за всеми достижениями компьютерной лингвистики. Вообще любое взаимодействие человека и компьютера возможна при наличии физического интерфейса (hardware) и программного интерфейса (software). И если за физическую сторону революции в компьютерной науке отвечали транзисторы, именно интеллектуальная схема отвечает за программный сторону взаимодействия.

Еще в 1950-е годы была выведена концепцию, согласно которой структура интеллектуальной компьютерной системы включает в себя три основных блока.

База знаний представлена в виде физического накопителя, на котором содержится в цифровом виде, пригодном для взаимодействия, информация, которой будет оперировать программное обеспечение.

Логический блок, или решатель, имеет основную функцию поиск вывода, ответы на входящий запрос, попадающий в систему. Если в базе знаний (базе данных) содержится информация, соответствующая поставленному вопросу или запросу, сформированному естественным языком, система коммуникации превращает естественный язык в его внутренний цифровой аналог, с которым и работает логический блок, расшифровывая его как запрос на доступ к базе знаний. В случае, если логический блок не находит прямого ответа на поставленный запрос, он за встроенной функцией устраивает поиск косвенной информации, которая так или иначе касается запроса.

Интеллектуальный интерфейс обеспечивает взаимодействие пользователя с компьютером и корректное функционирование компонентов интеллектуальной компьютерной схемы. Частью, а ныне основой и ядром интеллектуального интерфейса многих приложений есть искусственный интеллект.

Искусственным интеллектом называют свойства компьютерных интеллектуальных систем прибегать к творческим функциям, которые традиционно считаются присущими только человеку. Современные исследователи разделяют ИИ на сильный и слабый. Сильный представляет собой много ИИ, интегрированных в единую систему, достаточно мощную, чтобы решать даже общечеловеческие проблемы. Слабый – это узкоспециализированный ИИ, интегрированный с целью помогать человеку в определенных конкретных областях. В компьютерной лингвистике используется концепт слабого ИИ.

Среди мировых лидеров в разработке систем искусственного интеллекта можно назвать исследовательские центры в США – Массачусетский технологический институт (Massachussets Institute of Technology, MIT), Исследовательский институт машинного интеллекта (Machine Intelligence Research Institute, MIRI), – Японии – Национальный институт современной промышленной науки и технологии (AIST), а также Германии – Немецкий исследовательский центр по вопросам искусственного интеллекта (Deutsches Forschungszentrum für Künstliche Intelligenz, DFKI).

Самым полезным среди свойств искусственного интеллекта является то, что он способен к

обучению и самообучению. Над этим работают сразу несколько направлений на перекрестке информационных технологий и лингвистики. Схема машинного обучения – это искусственный интеллект, первостепенной направленностью которого является не прямое решение поставленной задачи, а обучение в процессе его применения, как решать много похожих задач. Различают два метода обучения ИИ: обучение по прецедентам и дедуктивное обучение.

Обучение по прецедентам, основанное на анализе компьютером конечной совокупности факторов – прецедентов – на основе обучающей выборки, на основе которых ИИ должен построить алгоритм, выдающий достаточно правильный ответ на каждый из объектов. Для измерения правильности ответов в ИИ встроен также функционал качества.

Дедуктивное обучение является прерогативой так называемых экспертных систем, и является следующим этапом – когда машина уже обучена на прецедентах, выполняет роль эксперта, анализирует имеющиеся данные на основе запроса и выдает экспертное заключение.

Итак, прояснив то, разработка чего именно способствовало возникновению компьютерной лингвистики, переходим к тому, в каких отраслях лингвистики компьютер помогает человеку, и каким образом.

Предоставленная выше схема в упрощенном виде помогает раскрыть спектр функций, с помощью которых компьютер помогает современным лингвистам.

Функция распознавания и оцифровки в широком смысле понимание – это способность компьютера с помощью интерфейса, физических сенсорных устройств или программных аналитических средств распознавать информацию, например:

– Преобразования речевого сигнала в цифровую информацию. Первое устройство для распознавания речи появилось в 1952 году, он мог распознавать цифры, которые произносил пользователь. Коммерчески использованные программы по распознаванию речи появились в 1990-х, частично они были призваны облегчить пользование компьютером для тех, кто не мог пользоваться клавиатурой как средством ввода информации вследствие перенесенных травм или ограниченных возможностей. Среди этих программ можно назвать Dragon NaturallySpeaking, VoiceNavigator, Microsoft Voice Command и другие. Механизм работы такой: обработка речи начинается с оценки качества речевого сигнала, определяется уровень помех и искажений. Результат оценки предоставляется к модулю акустической адаптации, который управляет модулем расчета параметров, необходимых для распознавания. Далее в сигнале выделяются участки, содержащие речь, и оцениваются его параметры, выделяются фонетические и иные характеристики, проводится синтаксический, семантический и прагматический анализ. Наконец, проанализирована информация поступает в декодер, что сопоставляет входящий речевой поток с информацией, имеющейся в акустической и языковой базе данных, и выдает наиболее вероятную последовательность слов, которая и станет конечным результатом.

Оценка качества Акустическая Оценка Декодирования и входного адаптация лингвистических выдача сигнала параметров результата

– Перевод рукописного текста на естественном языке в цифровую информацию;

– Распознавание текста на естественном языке на изображениях различного цифрового формата;

– Распознавание жестов, проведенных в режиме реального времени на камеру (сенсорное устройство) или предварительно записанных в видеоформате.

По разным вариантам распознавания два последних этапа, изображенные в схеме для распознавания речи, остаются неизменными, тогда как два первых будут варьироваться в зависимости от типа информации, которую нужно распознать.

Следующей функцией является статистика и каталогизация. Перед компьютерными лингвистами встает вопрос, как формализовать информацию, что сейчас имеется в распознанном электронном виде. Этим занимается корпусная лингвистика – раздел языкознания, занимающийся разработкой, созданием и использованием текстовых корпусов, то есть совокупностей текстов, отобранных в соответствии с определенными принципами. Эти принципы каталогизации заложены в ИИ, что отвечает за статистическую обработку и каталогизацию текстов, поступающих в базу знаний. Необходимость корпусов текстовых данных объясняется тем, что это является представлением лингвистической информации в определенном конкретном контексте; за большого объема корпуса есть также большое наличие информации для анализа; наконец, единожды созданный корпус имеет многоразовое применение для решения разнообразных лингвистических записей и проблем. На языке информационных технологий, корпусная лингвистика занимается созданием и укладкой хорошо структурированной базы данных, что касается языка в ее конкретных проявлениях – устных и письменных актах коммуникации. Первым созданным лингвистическим корпусом был так называемый корпус Брауна, созданный в начале 1960-х годов в университете Брауна. Он содержал 500 фрагментов, каждый из которых состоял из 2000 слов английского языка. Этот стандарт размера корпуса слов языка – 1 000 000 – был использован при создании аналогичных корпусов в 80-е годы XX века для русского, немецкого и французского языков. Сейчас в интернете существующий сайт Татоеба, что имеет целью на свободной основе добавлять и изменять предложения и их переводы, связанные между собой по смыслу, то есть объединять корпуса двух языков в единое переводческое пространство. Этот ресурс может пригодиться как компьютерным лингвистам, так и профессиональным переводчикам, так как настройки корпусов пары языков для отыскания четких эквивалентов и является одной из основ работы машинного перевода. Сейчас количество языков на ресурсе более 80, а количество предложений превысило отмутку в 600 000. Имеется также функция бесплатной загрузки всех корпусов – полностью или частично.

Аналитика и выводы. Существует общее направление развития ИИ на началах математической лингвистики, называемое обработкой естественного языка (англ. NLP, Natural Language Processing). Эта дисциплина изучает проблемы компьютерного анализа и синтеза естественных языков. В контексте ИИ анализ означает понимание языка, а синтез – генерацию на основе выводов качественного текста. Понимание естественного языка зависит от чрезвычайно большого количества факторов – лингвистических, экстралингвистических (культурных, социологических, исторических), личности собеседника и под. Среди лингвистических сложностей, с которыми сталкивается ИИ на этапе понимания текста, можно привести такие:

– Раскрытие анафор (то есть понимание того, что подразумевается при условии применения местоимений). Например, есть два предложения: «Мы удалили данные сотрудников, потому что они были повреждены» и «Мы удалили данные сотрудников потому, что они были скомпрометированы». В первом случае местоимение «они» касается слова «данные», во втором – слова «сотрудники». Правильное понимание компьютером смысла предложения зависит от того, знает ли он, какая характеристика подходит к какому из существительных.

– Свободный порядок слов языка может привести к неправильному пониманию фразы: «Порядок определяет процесс» – непонятно, что определяется чем, или порядок процессом, или же наоборот, процесс – порядку.

– Наличие в тексте неологизмов типа «смснуть» и «зафрендить», которые ИИ должна

отличать от орфографических и печатных ошибок и понимать.

– Омонимы типа «ключ\ключ», «зАмок\замОк» и проч. составляют проблему, поскольку требуют детального анализа и понимания речевого окружения, то есть контекста, прежде чем использовать (или переводить) их.

Среди имеющихся систем для анализа и обработки естественного текста можно назвать AlchemyAPI, Nautral Language Toolkit, MontyLingua, General Architecture for Text Engineering (GATE).

Синтез и перекодировки. Под синтезом в данном контексте понимаем генерацию качественного текста компьютером, а под перекодировкой – дополнительный этап перевода данного текста на другой язык перед этапом его финальной генерации. Именно эта функция компьютерных интеллектуальных схем с применением ИИ и подарила нам программы для машинного и автоматизированного перевода, на чем подробнее сосредоточимся ниже. Полный синтез речи по всем правилам (или синтез по печатному тексту) обеспечивает управление всеми параметрами речевого сигнала и, таким образом, может генерировать текст по ранее неизвестным текстам. Синтез реализуется путем моделирования и применения аналоговой или цифровой техники. Так, синтез устной речи требует имитации звуков речи в аналоговом или цифровом формате, а генерация компьютером письменного текста требует достаточного понимания принципов лингвистической организации текста, узусу, контекстуального словоупотребления и под.

Перекодирование же текста является аспектом компьютерной лингвистики, что касается непосредственно перевода, и поэтому представляет особый интерес для переводчиков. Вариантов компьютерного перевода текста, как уже отмечалось, два – машинный и автоматизированный.

Машинный перевод – процесс перевода текстов, поданных в письменном или устном виде, с одного естественного языка на другой с помощью специальной компьютерной программы. Существует четыре вида машинного перевода.

В условиях постредактирования исходный текст полностью перекодируется машиной, а человек-редактор корректирует результаты. По условиям предредактирования человек первично приспосабливает текст к обработке его машиной (убирает возможные неоднозначности, упрощает и структурирует текст), после чего начинается программка обработка. Интерредактирование предполагает, что человек вмешивается в процесс перевода, решая сложные случаи и проблемные вопросы. Смешанные системы означают комбинирование изложенных выше вариантов.

Вообще мысли о том, что можно использовать ЭВМ для решения вопросов перевода была высказана еще в 1947 году. Еще за 7 лет, в 1954, произошел так называемый Джорджтаунский эксперимент, известный как первая публичная попытка демонстрации машинного перевода. Интеллектуальная компьютерная система, которая использовалась тогда, содержала словарь из 250 слов, грамматика из 6 правил и базу знаний перевода из нескольких простых фраз. Несмотря на ее простоту, она вызвала настоящий бум в лингвистическом мире и стимулировала производные разработки во многих развитых странах. Уже в середине 1960х годов американскими разработчиками было представлено две системы машинного перевода с русского языка (несмотря на тогдашние актуальные обстоятельства холодной войны и необходимости разведки это было более чем логичным): MARC (разработка технического департамента BBC) и GAT (разработана учеными-исследователями Джорджтаунского университета). Однако, учитывая низкое качество переводов, финансирования и стимуляция

этих проектов были не слишком щедрыми, что притормозило развитие данной отрасли.

Качество машинного перевода зависит от самого программного обеспечения, но также и от тематики, стилистики, грамматики, синтаксиса и лексики текста, родства языков оригинала и перевода и других факторов. Научно доказано практичным образом, что, хотя художественный перевод вследствие многослойности текста и контекста не поддается качественному машинному переводу, некоторые типы технических или официально-деловых текстов получили машинный перевод приемлемого качества, требовал минимального вмешательства редактора\корректора.

Среди известных программ, использованных для машинного перевода, следует назвать Trados, PROMT, Multitran, SmartCAT, Google Translate, DejaVu X3, а также ЯндексПереводчик.

В любом случае степень вовлечения человека в процесс машинного перевода минимальный. Ситуация же с автоматизированным переводом другие.

Автоматизированный перевод – перевод текстов человеком на компьютере с применением компьютерных технологий. От машинного перевода он, собственно, отличается тем, что процесс перевода полностью осуществлен человеком, компьютер привлекается лишь к тому, чтобы ускорить процесс или улучшить его качество.

Интересно, что идея автоматизированного перевода в сыром виде была высказана еще в 1933 году, когда советский ученый П.П. Троянский изложил идею разработки машины для подбора и печати слов при переводе с одного языка на другой. Устройство состояло из стола с наклонной поверхностью, перед которым был закреплен фотоаппарат, синхронизированный с пишущей машинкой. На поверхности стола находилось «поле глоссария» – свободно подвижная пластина с напечатанными словами тремя, четырьмя и более языками. Понятно, что на тот момент идея была технически не реализована.

Заключение

На данный момент самыми распространенными способами применения интеллектуальных компьютерных систем для осуществления автоматизированного письменного перевода работа со словарями, глоссариями, памятью переводов и корпусами. Кроме того, есть специализированное программное обеспечение, применяемое для решения конкретных задач, например, аудиовизуального перевода, локализации продуктов программного обеспечения, редактирование текстов.

Библиография

1. Boguraev B. Machine-Readable Dictionaries and Computational Linguistics Research // Current Issues in Computational Linguistics: In Honor of Don Walker. Dordrecht: Springer Netherlands, 1994. P. 119-154. https://doi.org/10.1007/978-0-585-35958-8_7
2. Coscoy Y. A natural language explanation for formal proofs // Logical Aspects of Computational Linguistics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997. P. 149-167.
3. Dahl V., Maharshak E. DNA Replication as a Model for Computational Linguistics // Methods and Models in Artificial and Natural Computation. A Homage to Professor Mira's Scientific Legacy. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. P. 346-355.
4. Houben J. Paṇini's Grammar and Its Computerization: A Construction Grammar Approach // Sanskrit Computational Linguistics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. P. 6-25.
5. Huet G. Zen and the Art of Symbolic Computing: Light and Fast Applicative Algorithms for Computational Linguistics // Practical Aspects of Declarative Languages. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. P. 17-18.
6. Ide N., & Walker D. Introduction: Common methodologies in humanities computing and computational linguistics // Computers and the Humanities. 1992. 26(5). P. 327-330. https://doi.org/10.1007/BF00136978

7. Lecomte A. Categorial Minimalism // Logical Aspects of Computational Linguistics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001. P. 143-158.
8. McDonald D.D. On the Place of Words In the Generation Process // Natural Language Generation in Artificial Intelligence and Computational Linguistics. Boston, MA: Springer US, 1991. P. 229-247. https://doi.org/10.1007/978-1-4757-5945-7_9
9. Schmidt K.M. Concepts and Grammar: Thoughts about an Integrated System // Language, Culture, Computation. Computational Linguistics and Linguistics: Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. P. 14-35. https://doi.org/10.1007/978-3-642-45327-4_2
10. Shao Y. et al. Construction of Multilingual Terminology Bank of Computational Linguistics // Chinese Lexical Semantics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. P. 572-580.
11. Simon I. Sequence comparison: Some theory and some practice // Electronic Dictionaries and Automata in Computational Linguistics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989. P. 79-92.
12. Wu Y., Zhao H. Finding Better Subword Segmentation for Neural Machine Translation // Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Cham: Springer International Publishing, 2018. P. 53-64.

Formation of the goal of introducing computer linguistics for foresight purposes in economic development

Andrei I. Kapitanov

Assistant,
National Research University of Electronic Technology,
124498, 1, Shokina square, Zelenograd, Russian Federation;
e-mail: andrey@kapdx.ru

Abstract

None of the sources studied in this paper examines the role played by computer linguistics for translators, nor does it contain information useful for translators, issued the most important for involvement in practical translation activities. This determines the relevance of this article. Computational linguistics is a discipline required for study at the philological faculties, and it is for translators that it is of much greater interest, since they make it possible to translate faster and with higher quality, so the translator must be aware of the IT linguistic aspects of not one language, like a philologist, but two at once, which significantly increases the load. The problem of this article is to highlight why computer linguistics is of particular interest to beginners, professional translators and students of the specialty "Translation", as this will help the student to more effectively structure the learning process, and the specialist to supplement the gaps in existing knowledge. At the moment, the most common ways of using intelligent computer systems to carry out automated translation are working with dictionaries, glossaries, translation memory and cases. In addition, there is specialized software used to solve specific problems, for example, audiovisual translation, localization of software products, text editing.

For citation

Kapitanov A.I. (2019) Formirovanie tseli vnedreniya komp'yuternoi lingvistiki dlya tselei forsaita v ekonomicheskom razvitii [Formation of the goal of introducing computer linguistics for foresight purposes in economic development]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 9 (7A), pp. 23-31.

Keywords

Linguistics, economic development, foresight, IT, implementation.

References

1. Boguraev B. (1994) Machine-Readable Dictionaries and Computational Linguistics Research. In: *Current Issues in Computational Linguistics: In Honor of Don Walker*. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-0-585-35958-8_7
2. Coscoy Y. (1997) A natural language explanation for formal proofs. In: *Logical Aspects of Computational Linguistics*. Berlin, Heidelberg: Springer Berlin Heidelberg.
3. Dahl V., Maharshak E. (2009) DNA Replication as a Model for Computational Linguistics. In: *Methods and Models in Artificial and Natural Computation. A Homage to Professor Mira's Scientific Legacy*. Berlin, Heidelberg: Springer Berlin Heidelberg.
4. Houben J. (2009) Paṇini's Grammar and Its Computerization: A Construction Grammar Approach. In: *Sanskrit Computational Linguistics*. Berlin, Heidelberg: Springer Berlin Heidelberg.
5. Huet G. (2003) Zen and the Art of Symbolic Computing: Light and Fast Applicative Algorithms for Computational Linguistics. In: *Practical Aspects of Declarative Languages*. Berlin, Heidelberg: Springer Berlin Heidelberg.
6. Ide N., Walker D. (1992) Introduction: Common methodologies in humanities computing and computational linguistics. *Computers and the Humanities*, 26(5), pp. 327-330. https://doi.org/10.1007/BF00136978
7. Lecomte A. (2001) Categorial Minimalism. In: *Logical Aspects of Computational Linguistics.* Berlin, Heidelberg: Springer Berlin Heidelberg.
8. McDonald D.D. (1991) On the Place of Words In the Generation Process. In: *Natural Language Generation in Artificial Intelligence and Computational Linguistics*. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4757-5945-7_9
9. Schmidt K.M. (2014) Concepts and Grammar: Thoughts about an Integrated System. *In: Language, Culture, Computation. Computational Linguistics and Linguistics: Essays Dedicated to Yaacov Choueka on the Occasion of His 75th Birthday*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-45327-4_2
10. Shao Y. et al. (2013) Construction of Multilingual Terminology Bank of Computational Linguistics. In: *Chinese Lexical Semantics.* Berlin, Heidelberg: Springer Berlin Heidelberg.
11. Simon I. (1989) Sequence comparison: Some theory and some practice. In: *Electronic Dictionaries and Automata in Computational Linguistics.* Berlin, Heidelberg: Springer Berlin Heidelberg.
12. Wu Y., Zhao H. (2018) Finding Better Subword Segmentation for Neural Machine Translation. In: *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data.* Cham: Springer International Publishing.