

КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА ДЛЯ СТУДЕНТОВ

Карпіловська Є.А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика: Підручник. — Донецьк: ТОВ “Юго-Восток, Лтд”, 2006. — 188 с.

К. М. Шилихина

Воронежский государственный университет

В наших школах нет предмета под названием “лингвистика”, поэтому выпускникам школы, как правило, неизвестно, чем именно занимаются лингвисты. В результате для тех, кто выбирает лингвистику в качестве профессии, любая информация о практическом применении знаний о языке, как правило, оказывается новой. К тому же результаты практической деятельности лингвистов часто остаются незаметными для тех, кто не занимается языковыми проблемами на профессиональном уровне. Неудивительно, что у рядовых носителей языка периодически возникает вопрос о пользе лингвистических исследований. Однако ответить на вопрос о том, зачем надо исследовать язык, сегодня достаточно несложно, и вот почему. Если раньше пользу от деятельности лингвистов осознавали преимущественно те, кто изучал иностранные языки либо пользовался словарями или грамматиками (а в эпоху “железного занавеса” это была достаточно небольшая часть нашего общества), то сегодня круг людей, которые сталкиваются с результатами прикладной лингвистической деятельности в повседневной жизни, оказывается гораздо шире. Речь идет о тех, кто так или иначе использует компьютер. Именно появление компьютеров в нашей жизни задало новое направление лингвистических исследований, и сегодня вряд ли надо доказывать, что лингвистическая деятельность именно в области информационных технологий является остро необходимой.

История прикладной лингвистики насчитывает чуть более пятидесяти лет: традиционно ее появление связывают с Джорджтаунским экспериментом, начатым в 1949 году, однако до сих пор это научное направление считается молодым. Такое восприятие науки как новой объясняется целым рядом особенностей, которые в разной степени характерны для всех вновь созданных научных направлений. В первую очередь исследователи спорят о том, какие лингвистические исследования следует относить к кругу прикладных: следует ли

ограничиться только теми исследованиями, которые связаны с использованием лингвистических знаний в области новых информационных технологий, или любое практическое применение знаний о языке можно считать прикладным исследованием. Кроме того, ведутся споры относительно терминологии прикладной лингвистики, ее места в кругу других наук, связанных с проблемами обработки информации. В результате студентам, избравшим прикладную лингвистику в качестве своей будущей специальности, нелегко ориентироваться в потоке публикаций, так как сами лингвисты пока не могут четко очертить границы этой отрасли, поскольку направления прикладной лингвистической деятельности весьма разнородны.

Наиболее надежным ориентиром для начинающих изучать лингвистику остаются учебники и учебные пособия. За последнее десятилетие таких изданий на русском языке вышло несколько — учебник, созданный авторским коллективом под руководством А. С. Герда [1], книга Ю. Н. Марчука [2], учебники А. Н. Баранова [3], Р. К. Потаповой [4], А. В. Зубова и И. И. Зубовой [5]. В названных учебниках реализуются различные подходы к пониманию границ прикладной лингвистики: от включения в сферу ее интересов наиболее широкого круга проблем (такой подход представлен в работах [1] и [3]), до подробного анализа способов применения лингвистических знаний только в современных информационных технологиях (см. работы [2], [4], [5]). Такой подход к сужению проблематики не должен вызывать удивления: во-первых, как мы уже сказали, границы прикладной лингвистики до сих пор остаются предметом дискуссий, а во-вторых, компьютерная лингвистика — это один из аспектов прикладной лингвистической деятельности, который сегодня развивается наиболее активно, и поэтому заслуживает особого внимания.

В 2006 г. в Донецке вышла книга Е. А. Карпиловской “Вступ до прикладної лінгвістики: комп'ютерна лінгвістика”. Это учебное пособие, адресованное украинским студентам, начинающим

изучать прикладную лингвистику. Из названия становится ясно, что в пособии описывается только та часть проблематики прикладной лингвистики, которая связана с использованием компьютеров в лингвистических исследованиях.

Автор подробно анализирует содержание термина “компьютерная лингвистика”, четко определяет объект и предмет исследования нового научного направления. Объект исследования понимается максимально широко — в сферу интересов компьютерной лингвистики попадают и язык как система, и речь как реализация языковой системы в различных условиях коммуникации, и речевая деятельность как процесс.

Основной текст книги состоит из двух разделов. Первый раздел учебного пособия посвящен общим вопросам, связанным с существованием компьютерной лингвистики как самостоятельного научного направления: описанию места компьютерной лингвистики в кругу других наук, анализу метода лингвистического моделирования, способам представления лингвистической информации, а также обзору отдельных направлений прикладной лингвистической деятельности — компьютерной лексикографии и корпусной лингвистики.

Существуют различные точки зрения на то, какое место занимает компьютерная лингвистика в современной системе наук. Обобщая эти точки зрения, Е. А. Карпиловская наглядно показывает место компьютерной лингвистики в кругу других лингвистических дисциплин и связь компьютерной лингвистики с нелингвистическими науками. При этом акцент делается на прикладном характере нового научного направления: автор показывает, как знания, накопленные в теоретической лингвистике, могут применяться в различных системах автоматической обработки текстов.

Особое внимание Е. А. Карпиловская уделяет методологии прикладных лингвистических исследований в сфере информационных технологий. Основным методом компьютерной лингвистики — это моделирование лингвистических объектов. Проблема моделирования в языкознании на современном этапе стоит достаточно остро, поскольку язык — это очень сложный объект исследования. Поэтому модель в лингвистике трактуется неоднозначно (достаточно сказать, что еще в 1968 году А. Ф. Лосев выделил 28 способов трактовки и использования этого термина в языкознании). Для тех, кто сегодня начинает изучать компьютерную лингвистику, необходимо четкое понимание сути моделирования как метода научного познания,

осознание проблем, с которыми сталкиваются исследователи, используя моделирование в прикладных лингвистических разработках. Поэтому в рецензируемом учебном пособии дается достаточно полный обзор основной проблематики лингвистического моделирования, приводится классификация лингвистических моделей, четко формулируются требования, предъявляемые к моделям.

Поскольку целевой аудиторией книги являются студенты, автор структурирует текст таким образом, чтобы читатель мог сосредоточить свое внимание на терминологии, которую используют в компьютерной лингвистике. В конце каждой главы дается краткий глоссарий терминов и их определений, что является одним из достоинств учебника.

Подводя читателя к рассмотрению конкретных направлений деятельности по автоматической обработке языка, Е. А. Карпиловская посвящает третий параграф первого раздела базам данных и базам знаний, которые сегодня активно используются для организации лингвистической информации. Однако для успешной работы знания способов организации информации недостаточно. Будущий лингвист должен иметь представление о том, как компьютер может обработать языковую информацию. Поэтому дальнейшее изложение основ компьютерной лингвистики требует введение новых понятий — лингвистического алгоритма и лингвистического процессора.

Возможности практического применения теоретических знаний автор демонстрирует на материале конкретных разработок украинских исследователей в области компьютерной лексикографии и корпусной лингвистики. Эти направления исследований компьютерной лингвистики представляют особый интерес, поскольку результатами исследований могут пользоваться не только сами лингвисты, но и рядовые пользователи компьютеров. Описанию способов организации словарной информации в автоматических словарях разных типов посвящен пятый параграф первого раздела. Практике создания корпусов текстов посвящен следующий, шестой параграф.

Первый раздел учебника завершается описанием компьютерного фонда украинского языка, созданного в Институте языкознания им. А. А. Потебни.

Второй раздел посвящен рассмотрению проблем применения лингвистических знаний при создании интеллектуальных компьютерных систем. Первый и второй параграфы раздела посвяще-

ны анализу свойств человеческого и искусственного интеллекта, рассмотрению основных проблем моделирования человеческого интеллекта с помощью компьютера.

Важным компонентом системы искусственного интеллекта являются базы знаний, поэтому Е. А. Карпиловская знакомит читателей с различными способами организации знаний, а также со структурами представления различных типов знаний.

Необходимым условием успешного моделирования интеллекта является этап анализа текста на различных языковых уровнях. В учебном пособии Е. А. Карпиловской даются основы автоматического анализа текста на морфологическом, синтаксическом и логико-семантическом уровне.

Практическое применение подобные разработки находят в системах машинного перевода, в компьютерных диалоговых системах. Их описанию посвящены два последних параграфа второго раздела. В качестве примеров практической реализации таких систем автор приводит системы французско-русского перевода ЭТАП-1, англо-русского перевода ЭТАП-2, русско-украинского перевода ПЛАЙ.

Несомненным достоинством учебного пособия является его ориентация на конкретные разработки, выполненные на материале русского и украинского языков, обобщение опыта украинских и российских исследователей. Современное состояние литературы по данному вопросу таково, что студенту не всегда легко самостоятельно разобраться в тонкостях систем, особенностях терминологии, достоинствах и недостатках каждой конкретной разработки. Учебное пособие Е. А. Карпиловской предоставляет начинающему лингвисту базовую информацию о возможностях имеющихся систем анализа и синтеза текстов, морфологических анализаторов, систем машинного перевода. Автор не обходит стороной и проблемы, с которыми столкнулись лингвисты в процессе создания различных

программных продуктов. Тем самым намечаются перспективы дальнейших прикладных лингвистических исследований.

В конце учебного пособия автор дает краткий список терминов, который имеет алфавитно-иерархическую организацию. Это облегчает поиск необходимой терминологии, а также наглядно демонстрирует читателям логические отношения между терминами.

Подводя итог, можно сказать, что данное учебное пособие является надежным проводником в мир компьютерной лингвистики. Книга может стать опорой для тех, кто в силу своих профессиональных интересов должен владеть методами компьютерной обработки лингвистических данных. Работая с этой книгой, студенты имеют возможность получить представление о круге проблем, которыми интересуется современная прикладная лингвистика, приобрести базовые знания в этой области, овладеть терминологией и сформировать представления о дальнейших перспективах развития этого направления.

ЛИТЕРАТУРА

1. Прикладное языкознание: Учебник / Авт. С. А. Аверина, И. В. Азарова, Е. Л. Алексеева и др. Отв. ред. А. С. Герд; С. Петербург. гос. ун-т. — СПб.: Изд-во С-Петербург. ун-та, 1996. — 525 с.
2. Марчук Ю.Н. Основы компьютерной лингвистики: учеб. пособие / Ю. Н. Марчук; Моск. пед. ун-т. — М.: Народный учитель, 2000. — 227 с.
3. Баранов А.Н. Введение в прикладную лингвистику / А. Н. Баранов; Моск. гос. ун-т им. М. В. Ломоносова. Фил. фак. — М.: Эдиториал УРСС, 2001. — 358 с.
4. Потапова Р.К. Новые информационные технологии и лингвистика: учебное пособие для студ. вузов / Р. К. Потапова; Моск. гос. лингв. ун-т. — Изд. 2-е. — М.: Эдиториал УРСС, 2004. — 317 с.
5. Зубов А.В. Информационные технологии в лингвистике: учебное пособие для студентов вузов, обучающихся по специальности 021800 — Теоретическая и прикладная лингвистика / А. В. Зубов, И. И. Зубова. — М.: Academia, 2004. — 205 с.