

ПЕДАГОГИКА И ФИЛОЛОГИЯ

ТЕМАТИКА И ОБЪЕМ СЛОВАРНЫХ БАЗ ЭЛЕКТРОННЫХ СЛОВАРЕЙ

Г. Г. Бабалова, доцент кафедры иностранных языков Омского юридического института, кандидат филологических наук, доцент

Одним из критериев оценки электронных словарей является их объем. В бумажном словаре принято говорить о «количестве слов», т. е. словарных статей. Для электронного словаря определить объем и сравнить его сложнее. Словарь может представлять собой сложную структуру индексов поисковых терминов и связей со словарными статьями. Такое усложнение позволяет делать поиск намного более быстрым и эффективным. В настоящее время на рынке программной продукции наиболее популярными являются такие электронные словари, как ЛИНГВО 8.0 (Компания АВВУУ Software House (B T Software))¹, МУЛЬТИЛЕКС 3.5 (Компания «МедиаЛингва»)², КОНТЕКСТ 4.0 (Компания АО «Информатик»)³, ПОЛИГЛОССУМ 3.2 (Компания ETS Publishing House)⁴. Все рассматриваемые словари обеспечивают возможность расширенного поиска, т. е. поиска не только среди заглавных слов, но и внутри словарных статей. Таким образом, можно находить словосочетания, примеры, синонимы и даже слова, отсутствующие в основном массиве заглавных слов.

Все четыре коллекции словарей используют совершенно разные подходы и методы представления словарных данных. В результате объемы включенных в них словарей указаны разными способами: в словарных статьях, поисковых терминах или «единицах». «Статья» соответствует обыкновенной словарной статье при алфавитном поиске (как в бумажном словаре). Однако количество словарных статей как показатель объема – величина весьма обманчивая. На практике значимым оказывается то количество возможных вопросов, на которые в базе данных словаря содержатся точные ответы. Поэтому издательство ЭТС использует другую единицу измерения – «поисковый термин». Количество терминов в словаре равно количеству уникальных вхождений в словарь. Устойчивые сочетания, идиомы и т. п. считаются единым целым, а варианты перевода – отдельными вхождениями. Подобную систему применяет Контекст.

Словарные библиотеки различных словарей имеют свои особенности. Словари МультиЛекс стремятся быть лишь усовершенствованными ко-

¹ См.: Электронный словарь Контекст 4.0. – М.: Информатик, 2001.

² См.: Электронный словарь Лингво 8.0. – М.: Изд-во Бит, 2002.

³ См.: Электронный словарь Мультилекс 3.5. – М.: Медиа, 2001.

⁴ См.: Электронный словарь Полиглоссум 3.2. – М.: ETS Publishing House, 2001.

пиями своих печатных оригиналов. В них последовательно воспроизводится содержание печатных словарей, включая предисловия авторов оригинальных изданий и сопутствующие справочные материалы. Все словари МультиЛекс однонаправленные, а в качестве объема словаря указан объем оригиналов. Для обеспечения возможности перевода в обоих направлениях в МультиЛекс должна быть представлена пара независимых взаимодополняющих словарей. Например, «Новый Большой англо-русский словарь» (НБАРС) Ю. Д. Апресяна и «Русско-английский словарь» А. И. Смирницкого составляют такую пару общелексических словарей в выпуске МультиЛекс 4.0 «Для профессионалов».

Словари Lingvo не являются точными копиями бумажных аналогов. Так, в качестве источников для экономического словаря LingvoEconomics указаны семь словарей различных авторов. Некоторые словари из многоязычного набора Lingvo 8.0 сформированы и дополнены лексикографическим отделом фирмы АБВУУ, некоторые лицензированы у таких издательств печатных словарей, как «Русский язык» и «Руссо». Все словари Lingvo, кроме специальных («Грамматический словарь» и «Словарь произношения»), являются двунаправленными с точки зрения пользователя. При этом часть словарей представляет собой взаимодополняющие пары, как в МультиЛекс, а большинство просто снабжено системой индексов (ссылок), позволяющих вести обратный поиск в однонаправленном, в сущности, словаре.

Базы данных в Polyglossum формируются из единого источника с выпускаемыми издательством бумажными словарями и снабжены развитой системой индексации данных, позволяющей вести полнотекстовый поиск в любом направлении. Все словари Polyglossum двунаправленные. Их объем указывается в «терминах», что учитывает особенности и новые возможности электронных версий словарей. Среди словарей Polyglossum есть авторские бумажные словари, переведенные в электронный формат без права внесения изменений. Имеются бумажные словари, в отношении которых издательство ЭТС имеет права на дальнейшее развитие (обновление, внесение поправок и толкований), а также издание в любой форме. Кроме того, есть электронные словари, не имеющие бумажных аналогов, такие как «Нефть - Газ - Химия», «Словарь по легкой промышленности».

Электронные словари Контекст подготавливаются и пополняются лексикографическим отделом компании «Информатик» по материалам печатных и авторских словарей. В электронном виде словари Контекст двусторонние, причем статьи в словарях симметричны в плане поисковых возможностей. Основу каждой статьи составляет множество синонимичных слов или словосочетаний

на двух языках. Каждый из синонимов рассматривается как поисковая единица и может быть снабжен пометой (часть речи, транскрипция и т. п.). Статья может содержать примеры переводов фраз, также рассматриваемых в качестве поисковых единиц. Часть статей может быть снабжена толкованием. При подготовке электронного словаря на основе бумажного аналога такая переработка оказывается довольно трудоемкой, но позволяет структурировать информацию, выдаваемую по запросу, так, чтобы уменьшить объем не относящихся к делу сведений. Всего в составе Контекста насчитывается 48 словарей. Объем словарей подсчитывается в поисковых единицах, приходящихся на синонимы, и не включает в себя примеры, хотя по примерам также ведется поиск.

Lingvo, Polyglossum и МультиЛекс включают в себя много словарных баз разного назначения. Разобраться в этом книжном море нелегко, но возможно. В целом многие словари эквивалентны друг другу еще и потому, что имеют единую основу. Достаточно сравнить некоторые словари, чтобы увидеть множество совпадений.

Анализ данных вышеупомянутых словарей позволяет сделать вывод о том, что наиболее полным по объему и широким по охвату тематики является словарь Polyglossum. Общий объем словарных статей включает 3 442 000 терминов. В нем представлены даже такие узко-специальные темы, как библиотечное дело, авиация, космос, детали машин, геодезия и картография, экология, пищевая и текстильная промышленность и многие другие. На втором месте стоит словарь Контекст. Его объем составляет 2 312 000 единиц. Преимущество этого словаря в том, что в нем наиболее полно отражены такие темы, как вычислительная техника, геофизика, геология. Примерно равными по объему являются словари МультиЛекс и Lingvo: соответственно 1 222 700 и 1 217 500 словарных статей. Данные словари широко охватывают такие темы, как медицина, техника, полиграфия. Хорошо представлена здесь общеупотребительная лексика.

В последние десятилетия общий кризис коснулся и процесса создания словарей. Издатели стремятся искать новые пути пополнения словарных баз. Так, фирма АБВУУ начала распространять инструментарий для создания словарей в специальном «пользовательском» формате и поощрять создание и распространение таких словарей через Интернет-портал www.Blingvo.ru. Среди авторов словарей проводятся конкурсы на лучший словарь и лучший макрос для MS Word, облегчающий процесс создания словарей для Lingvo. Число таких словарей уже достигло 38.

Практически в то же время среди электронных словарей издательства ЭТС появился «Пользовательский пополняемый англо-французско-немец-

ко-финско-русский словарь» В. А. Бутакова. Этот словарь использует особую версию оболочки Polyglossum и устанавливается вместе со всем необходимым инструментарием для импорта словарей в текстовом виде и для их редактирования. Также имеется очень полезная функция экспорта пользовательских словарей в формате Rich Text (.rtf). В этом заложено коренное отличие от политики ABBYY, чей DSL компилятор из комплекта Lingvo осуществляет лишь одностороннее преобразование. Инструментарий ЭТС, напротив, позволяет преобразовать личный словарь обратно в текстовый вид, тем самым избегая привязки к редкому закрытому формату данных. Так же, как и ABBYY, ЭТС предоставляет возможность публикации своего мини-словника, но не отдельно, а в составе «Пользовательского словаря».

Контекст и Мультилекс также предоставляют возможность ввести свои собственные словари, но не предлагают отлаженного пути их последующей публикации. Можно отметить тот факт, что для хранения данных личных словарей используется открытый формат XML, который является стандартным средством обмена информацией между базами данных. Достоинством этого способа оказывается простота использования и переносимость. Эти словари можно просматривать посредством обычного интернет-навигатора. Словари пользователя старых версий Контекста могут быть преобразованы в новый формат. Контекст включает все необходимые средства для создания пользовательских словарей, которых может быть неограниченно много. Создание новой статьи происходит путем заполнения несложной экранной формы.

Можно предположить, что результаты работы над свободно распространяемыми пользовательскими словарями Lingvo и Polyglossum будут впоследствии использованы для пополнения соответствующих словарных баз данных, и предложенные пользователями словарные статьи, пройдя редактуру, окажутся доступны в продаваемых издательствами электронных словарях. К сожалению, многие специалисты и корпорации все еще находятся в плену представлений о необходимости хранения в строгой тайне результатов своего словарного труда и отказываются делиться накопленными данными с коллегами, тем самым вынуждая друг друга многократно «открывать Америку».

Все описанные словари в большой степени универсальны, но вместе с тем каждый из них тяготеет к определенной нише. Большая часть усилий издателей коммерческих электронных словарей сфокусирована на английском языке, получившем широкое распространение в России, в том числе с косвенной государственной поддержкой. В меньшей степени популярен немецкий язык, а вот выбор словарей испанского, итальянского и даже

французского языков можно назвать скудным. Огромной редкостью у нас оказываются электронные переводные словари языков Восточной Европы и остального мира.

Для систематически изучающих иностранный язык и преподавателей лучшим выбором окажется Мультилекс. Он представляет очень точные аналоги известных бумажных словарей, дополненных системой полнотекстового поиска, далеко опережающей печатные версии. Словарные статьи содержат значительное количество дополнительных сведений и ссылок на большое число примеров. Большим достоинством словаря с точки зрения преподавателя иностранного языка оказывается наличие транскрипции, проставленное в русских и английских словах ударение (НБАРС), стилевые пометы и подробность разработки значений. Словарь больше подходит для художественного перевода и обучения благодаря явному предпочтению общелексических словарей перед специальными и подчеркнутой книжности подачи материала. Мультилекс – естественный выбор для изучающих романские языки, так как другие коллекции просто не имеют итальянских и испанских общелексических словарей, хотя в области французского языка с Мультилексом может поспорить Lingvo.

Лучший рабочий инструмент профессионального технического переводчика, инженера, ученого – Polyglossum. Издательство ЭТС – безусловный лидер по количеству, разнообразию и точности отраслевых и технических словарей. Общее количество предлагаемых сегодня словарных баз данных превышает сотню, в том числе имеется более 50 словарей английского и более 30 словарей немецкого языков. Коньком издательства являются качественные словари финского языка. Кроме того, существует шведский политехнический словарь для Polyglossum. Спектр языков словарей ЭТС явно ориентирован на страны северной Европы, имеющие тесные связи с Россией. Пока сравнительно слабо представлены романские языки. Polyglossum является наименее удобным в представлении информации (отсутствует наглядность), однако он обладает рядом других преимуществ: минимум занимаемого места на жестком диске и совместимость со всеми версиями операционной системы Windows, начиная с 3.11.

Lingvo полезен там, где требуется быстрая справка «по ходу дела» благодаря своему исключительно удачному интерфейсу и широкому выбору словарей. В этой роли с ним мог бы соревноваться Контекст, но он менее известен и доступен широкой публике. Lingvo – самая популярная оболочка, используемая не только пользователями Windows благодаря Wine и другим средствам эмуляции. Комплект Lingvo совмещает простоту и скорость доступа к данным с интуитивностью и хоро-

шим качеством словарей. Он может оказаться ценным дополнением к Polyglossum за счет дополнительных аббревиатур, введенных в отраслевые словари, особенно если самые «свежие» словари сокращений Polyglossum недоступны.

Контекст очень удобен для иностранцев, изучающих русский язык, благодаря возможности распознавания вводимых слов в любой форме, для получения справки, не отвлекаясь от основного занятия, но основной его аудиторией считаются переводчики. С точки зрения перевода достоинства Контекста можно считать: достаточно большой набор тематических словарей, возможность поиска словосочетаний в любой форме, а также возможность показать список переводов всех фраз с интересующим словом. Программная обо-

лочка Контекста достаточно проста, но имеет ярко выраженное преимущество: после ввода слова сразу осуществляется его морфологический анализ. Такой «интеллект» программы позволяет ввести для поиска слово в любом падеже или склонении. Программа определит его основу, найдет нормальную форму и уже по ней – словарную статью.

Вполне естественно, что разработчикам еще есть над чем потрудиться: подлежат совершенствованию системы морфологического анализа слов, словарные базы, необходимо пополнять программы новыми словарями специализированных терминов. В целом же, электронные словари имеют все средства для удобной и быстрой работы, избавляя от нужды копаться в традиционных бумажных изданиях.

ОСОБЕННОСТИ СОДЕРЖАНИЯ РАЗДЕЛА «МАТЕМАТИКА» УЧЕБНОЙ ДИСЦИПЛИНЫ «ИНФОРМАТИКА И МАТЕМАТИКА» В ЮРИДИЧЕСКОМ ВУЗЕ

М. А. Екимова, старший преподаватель кафедры правовой информатики Омского юридического института, кандидат педагогических наук

Важной особенностью системы образования, значительно повышающей качество профессиональной подготовки специалиста, на сегодняшний день является его фундаментализация. Причем первостепенная задача состоит в преодолении разобщения и даже некоторого противопоставления естественнонаучного и гуманитарного компонентов культуры путем их взаимообогащения. Поэтому введение естественнонаучных дисциплин, в том числе математики, в систему подготовки специалистов гуманитарного профиля является обоснованным и закономерным шагом.

Учебная дисциплина «Информатика и математика», изучаемая в юридических вузах, состоит из двух совмещенных разделов: «информатика» и «математика». Логика такого соединения не совсем понятна, ведь и информатика, и математика являются самостоятельными науками, имеющими свою специфику, поэтому уместнее бы было изучать их как две отдельные дисциплины. Конечно, эти науки сильно связаны, и в процессе обучения необходимо осуществлять их межпредметные связи. Но математика не должна выполнять только роль ма-

тематического аппарата для информатики, ее цели и задачи значительно шире.

Включение математики в учебные планы вузов, готовящих будущих юристов, вызвало ряд методических проблем, связанных с уточнением целей обучения математике, отбором и конструированием содержания обучения, выбором форм, методов и средств обучения, позволяющих наиболее полно реализовать поставленные цели. На сегодняшний день нет общепринятого и научно обоснованного подхода к проектированию методической системы обучения математике будущих юристов. Существуют проблемы и с учебной литературой для студентов юридических специальностей. Учебников много, но в них предложены различные подходы к изучению математики, существенно отличается даже содержание дисциплины, мало задач, профессионально значимых для будущего юриста.

Имеющийся образовательный стандарт по математике для студентов специальности «юриспруденция» предусматривает ее изучение с развивающей и общеобразовательной целью. Он неоднозначен, ориентирован на общее знакомство