

УДК 81'374:32=161.1

Л. Е. Бессонова

ЛЕКСИКОГРАФИЧЕСКОЕ МОДЕЛИРОВАНИЕ ПОЛИТИЧЕСКОЙ ЛЕКСИКИ: ОПЫТ СОСТАВЛЕНИЯ ЭЛЕКТРОННОГО СЛОВАРЯ

Создание комплексных лексикографических систем становится сегодня все более совершенным и представляет собой самостоятельное направление в современной лексикографии. В последнее десятилетие лингвисты активно обсуждают такие вопросы, касающиеся теории электронных словарей, как их структура и объем, построение базы данных и принципы системной организации лексики, систематизация информационного материала входного слова и методы лингвистического программного обеспечения (А. Н. Баранов, А. С. Герд, Б. Ю. Городецкий, Ю. Н. Караулов, А. А. Поликарпов, В. А. Широков и др.).

Основные аспекты этих и многих других научных проектов связаны не только с теоретическим исследованием естественного языка с позиций системно-количественного анализа, но и, что очень важно, с поиском новых возможностей использования компьютерных технологий при описании языковых единиц. Это и обусловило актуальность обозначенной проблемы, связанной, на наш взгляд, с организацией гипертекстового пространства в лексикографической практике и построением коммуникативно-дискурсивной модели словаря политического языка.

Такое решение экспериментальных задач привело к новому направлению компьютерной лингвистики — корпусной лингвистике, которая, по словам В. А. Широкова, “ставит своей задачей создание компьютерного инструментария для проведения лингвистических исследований на неоднородных массивах языковых данных практически неограниченного объема” [11, с. 339]. Так, созданный еще в 1960-ые годы Уппсальский корпус русских текстов остается, насколько нам известно, единственным завершенным и активно используемым проектом такого рода [10]. Однако устарелость его материалов, ограниченность объема (1 миллион словоупотреблений), а также отсутствие аннотирования текстов требуют дополнительной работы, которая проводится сегодня, например, в Институте проблем передачи информации РАН [5], Украинском языково-информационном фонде Национальной академии наук Украины [12].

Опыт разработки корпусов текстов как новой технологии прикладной лингвистики достаточно полно описан в лингвистической литературе [3, с. 112-137; 7, с. 193-195]. В этом научном обзоре четко обозначается круг проблем, одна из которых связана с отсутствием единого формата представления корпусов текстов, а также с принципами их организации. Основной параметр организации корпуса — аннотация, или разметка, является главной характеристикой корпуса. Она-то и отличает его от коллекций, библиотек, энциклопедий, широ-

ко представленных в современном Интернете [4]. Как известно, чем богаче и разнообразнее разметка, тем выше научная и учебная ценность корпуса. В Национальном корпусе русского языка сегодня разработаны пять типов разметки: метатекстовая, морфологическая, акцентная, синтаксическая и семантическая. Так определяются важные параметры научного описания языка со структурно-семантической и функциональной позиций. Пространство электронных текстовых корпусов дает возможность результативного использования *электронных конкордансов*, которые открывают перспективы моделирования языковой картины мира. Традиционное понимание конкорданса как специализированной лингвистической прикладной программы помогает решить проблему, во-первых, формирования корпуса текстов и, во-вторых, его использования в конкретных задачах (например, при автоматической выборке заданных языковых единиц) [6, с. 72].

Безусловно, корпусная лингвистика ярко и четко репрезентирует себя в одном из перспективных направлений — *компьютерной лексикографии*. Электронные словари и энциклопедии разрабатываются сегодня как автономные и сетевые программные продукты. Не вызывает сомнения тот факт, что электронные словари предоставляют пользователю множество дополнительных возможностей по сравнению с печатными аналогами:

- хранение большого объема информации за счет гиперссылок: электронный словарь предполагает включение в свою структуру нескольких словарей разных типов и жанров (словаря сочетаемости, терминологии, грамматических норм и др.);
- доступность словаря за счет эффективной системы поиска (полнотекстовый поиск, одновременный поиск в нескольких словарях, скорость поиска);
- применение средств мультимедиа для семантизации лексики: озвучивание заголовочных слов, введение иллюстративного материала с фотографиями, анимацией, видеофрагментами;
- использование словарей в локальной и глобальной сетях, причем работа со словарями может проводиться многими пользователями одновременно;
- экономия времени и материальных затрат при создании компьютерных словарей.

Значительный опыт исследований в компьютерной лексикографии позволил сегодня расширить возможности описания смысловой структуры политического слова. В этом исследовательском контексте формулируется основная цель словаря — определение коммуникативного статуса политического слова и словосочетания. Не останавливаясь на уточнении и определении терминов “политическое слово”, “идеологема”, “советизм”, а также на обосновании их выбора, определим специфику макроструктуры электронного словаря политических слов.

1. Словарь представляет собой многоуровневое лексикографическое произведение, в котором нашли отражение языковые факты политической жизни конца XX – начала XXI века.

2. Словарь построен по гипертекстовой структуре и включает информацию в виде лексических баз данных реляционного типа, когда инфор-

мация о лексеме распределяется по полям, причем главное внимание уделяется системам представления информации, характерной для целого класса лексем.

3. Словарь организует гипертекстовое пространство, которое является отражением когнитивного пространства человека.

4. Словарь оформлен по алфавитному принципу и принадлежит к толково-справочным с элементами энциклопедического типа. Наряду с собственно лингвистической информацией в нем даются сведения по истории слова, его функционированию в различных сферах.

5. Ориентированность на словарное произведение малого объема определило включение в словарь единиц, принадлежащих ядру современного политического лексикона — актуальной и употребительной политической лексики и фразеологии. Состав и объем словника, как известно, является одним из важных вопросов, стоящих перед лексикографами. Поэтому при разработке электронного словаря не следует ограничиваться специальной политической терминологией, необходимо учитывать и тот коммуникативно-прагматический потенциал лексического значения, который выводит слово с идеологизированным компонентом в политический контекст.

6. Материал словаря обладает коммуникативной направленностью.

Эти параметры могут быть реализованы, на наш взгляд, в следующей структуре словаря, предлагаемой для обсуждения.

Слово в электронном варианте реализуется в семантическом макрополе, состоящем из пяти полей, выстроенных в определенную систему:

1. Поле индексации служит для удобного поиска единицы в базе данных, которая создается в системе управления Access.

2. Поле заголовочного слова характеризуется традиционной, исходной, формой подачи и включает в себя несколько блоков.

2.1. Блок помет о произношении слова или его форм: информация об особенностях произношения помещается непосредственно после входного слова и представляет собой указание на нормативное произношение слова или его допустимый вариант. Слово может быть озвучено средствами мультимедиа.

2.2. Блок грамматической информации, прагматическая нагрузка которой связана с конкретным грамматическим статусом слова в текстах. Блок содержит основные формы заголовочного слова, его грамматические характеристики, что соответствует правилам отечественной лексикографии.

2.3. Блок этимологических справок: слово в лексикографическом описании любого типа апеллирует не только к синхронии, но и к диахронии.

2.4. Блок толкования строится не по всей семантической парадигме слова, а лишь по ее основной части, тематически связанной с политической сферой употребления. Эта часть микроструктуры словаря сопровождается примерами, речениями.

2.5. Блок фразеологии.

3. Поле иллюстраций работает для верификации дефиниции, тем самым выявляет семантику политического слова в определенном контекстуальном

окружении. Важно отметить, что иллюстрации из политических текстов различных жанров (материалы Корпуса газетных текстов русского языка конца 20-го века [6], Интернет-ресурсов, СМИ) используются и в традиционном виде как оправдательный контекст (это особенно важно для метафорических значений) и как лексикографический прием, помогающий создать пространственную, временную и социально-историческую перспективы. Репрезентативный текстовый материал, организованный с помощью электронного конкорданса, позволяет моделировать политическую картину мира и ее составляющие.

4. Поле толковых словарей как микроструктура электронного словаря включает дефиниции толковых и толково-справочных словарей, основной материал которых составляет общественно-политическая лексика. Таким образом определяется репрезентативность лексической единицы и ее роль в формировании политического языкового пространства.

5. Поле энциклопедических сведений отражает основные этапы становления данной дефиниции, эволюцию его доминантных признаков, а также дает дополнительные сведения об истории слова.

При лексикографической реализации семантики слова раскрывается тот прагматический потенциал, который составляет эмпирическое поле словаря. Что же понимать под прагматической ценностью словарного материала? Вопрос, поставленный в лингвистической семантике в 80-90-х годах [1], до сих пор не получил однозначного решения, однако обозначил новый вектор в широкой научной дискуссии. В статье “Два принципа и два понятия системной лексикографии” (2005) Ю. Д. Апресян, развивая свою концепцию интегральности лексикографического описания лексемы, описанную в начале 90-х гг. [2, с. 270-271], выделяет семантико-прагматическое представление языковой единицы как особый тип лексикографической информации в словарных статьях интегрального словаря.

Такое широкое понимание прагматического компонента, а именно: отношение к действительности, к содержанию сообщения и к адресату — позволяет делать вывод о коммуникативно-прагматическом потенциале единиц лексикографического описания. Г. Н. Складневская, автор больших научных проектов, при лексикографировании слов выходит за рамки традиционных способов описания, выделяя в словарных статьях коммуникативно-ситуационный, когнитивный, эмотивный и идеологический компоненты как маркеры речевой ситуации. Как отмечает исследовательница, при формировании структурированной электронной картотеки целесообразно выделить особую прагматическую зону значения, которая позволит дать более полное и адекватное языковой действительности представление о содержании лексического значения и его структуре [9, с. 79]. Это справедливое замечание должно учитываться при систематизации и построении словника электронного словаря: *автократия, амбиция, бюрократия, вертикаль, вето, власть, глобализм, двоевластие, дежа вю, демарш, демократия, европеизация, импичмент, интеграция, клан, консенсус, лобби, неофит, олигарх, оппозиция, позиционирование, популизм, портфель, терроризм, толерантность, фигурант, электорат и т. д.*

Вопрос о прагматическом компоненте в словарной практике определен и в работах Л. П. Крысина, который предлагает систему прагматических помет: “Одни из этих помет сходны со стилистическими, другие указывают на определенные компоненты в семантике, третьи — на положительные или отрицательные коннотации слова, сопровождающие те или иные его значения” [8, с. 302].

Изложенные предварительные соображения дают основания для вывода о том, что семантическая информация при комплексном лексикографическом описании пересекает границы словарной статьи, представляя политическое слово в понятийно-концептуальном аспекте. Кроме того, данное словарное исследование дает возможность выявить основные тенденции в развитии ядерной политической лексики, обеспечить доступность словаря за счет эффективной системы поиска, а также хранить большой объем информации за счет гиперссылок.

Литература

1. Апресян Ю. Д. Прагматическая информация для толкового словаря // Апресян Ю. Д. Избранные труды. Т. 2. Интегральное описание языка и системная лексикография. — М., 1995. — С. 135-155.
2. Апресян Ю. Д. Два принципа и два понятия системной лексикографии // Язык. Личность. Текст: Сб. ст. к 70-летию Т. М. Николаевой / Ин-т славяноведения РАН. — М.: Языки славянских культур, 2005. — С. 267-303.
3. Баранов А. Н. Введение в прикладную лингвистику: Учебное пособие. Изд. 2-е, исправленное. — М.: Едиториал УРСС, 2003. — 360 с.
4. Библиотека Максима Мошкова <http://lib.ru/>. Русская виртуальная библиотека <http://www.rvb.ru/>. Фундаментальная электронная библиотека <http://febweb.ru/>.
5. Богуславский И. М. и др. Аннотированный корпус русских текстов: концепция, инструменты разметки, типы информации // Труды Международного семинара по компьютерной лингвистике и её приложениям «Диалог-2000». Протвино, 2000.
6. Герд А. С., Богданов В. В., Азарова И. В., Аверица С. А., Зубова Л. В. Автоматизация в лексикографии и словари-конкордансы // Филологические науки, 1981. — №1 (121). — С. 72-78.
7. Компьютерный корпус газетных текстов русского языка конца 20-го века <http://www.philol.msu.ru/~lex/korpus.html>.
8. Михайлов М. Н. Компьютерное обеспечение корпуса текстов (взгляд пользователя) // Русистика сегодня, 1998. — №1-2. — С. 192-201.
9. Крысин Л. П. О типах лексикографической информации в русской части русско-иноязычных словарей // Язык. Личность. Текст: Сб. ст. к 70-летию Т. М. Николаевой / Ин-т славяноведения РАН. — М.: Языки славянских культур, 2005. — С. 285-303.
10. Склярёвская Г. Н. Реальный и ирреальный мир в толковом словаре (к вопросу о прагматическом компоненте слова) // Семантика и коммуникация. — СПб: Изд-во СПб ун-та, 1996. — С. 68-80.

11. Уппсальский машинный фонд русского языка <http://www.slaviska.uu.se/korpus.html>. Британский национальный корпус (BNC) <http://sara.natcorp.ox.ac.uk/lookup.html>. Чешский национальный корпус <http://ucnk.ff.cuni.cz>.

12. Широков В. А. К основаниям теории лексикографических систем // Русский язык сегодня. Вып. 3. Сб. ст./РАН. Ин-т рус. яз. им. В. В. Виноградова. Отв. ред. Л. П. Крысин. — М., 2004. — С. 338-347.

13. Широков В. А., Манак В. В. Організація ресурсів національної словникової бази // Мовознавство, 2001. — №5.

Статья поступила в редакцию 9 ноября 2005 г.

УДК 811.161.1?243(038)

С. В. Варава

ОСОБЕННОСТИ ПОСТРОЕНИЯ СЛОВАРНОЙ СТАТЬИ В УЧЕБНОМ СЛОВАРЕ-СПРАВОЧНИКЕ “1201 РУССКИЙ ГЛАГОЛ”

Общая постановка проблемы. Существует немало хороших словарей-справочников, в которых можно найти справочную информацию о русских глаголах, которыми мы — преподаватели — пользуемся в своей работе. Но, с нашей точки зрения, они недостаточно удобны для работы с ними иностранных студентов на начальном этапе. Нам представлялось важным и целесообразным создание словаря-справочника глагольных форм, в котором достаточно наглядно были бы представлены все глагольные формы, а также модели синтаксической сочетаемости глаголов и с которым поэтому было бы удобно работать студентам с самых первых дней учебы на подготовительных факультетах.

Цель статьи. В настоящей статье мы попытаемся дать краткую характеристику некоторым уже существующим словарям глаголов и проанализировать особенности, отличающие построение словарной статьи, от словарей, с которыми мы традиционно привыкли работать.

Анализ последних исследований. Среди наиболее интересных словарей можно отметить словарь “Русский язык для всех” [1], который содержит около 35 000 слов современного русского литературного языка и представляет собой лексикографическое справочное издание нового типа, в основе построения которого лежит принцип комплексного описания слова как лексической, грамматической и словообразовательной единицы языка. Такой подход предопределил и емкое содержание словарной статьи, из которой иностранец, изучающий русский язык, узнает грамматическую информацию о слове, его стилистическую принадлежность, этимологическую справку, производные слова и устойчивые обороты, включающие заглавное слово. Важным элементом статьи яв-