

ЯЗЫКОЗНАНИЕ

УДК 81'44

К ИСТОРИИ ВОПРОСА О ФОРМАЛИЗАЦИИ СОПОСТАВИТЕЛЬНО-ТИПОЛОГИЧЕСКОГО ОПИСАНИЯ ЯЗЫКОВ

канд. филол. наук, доц. **Е.П. БЕТЕНЯ**
(Минский государственный лингвистический университет)
elbet@mail.ru

В статье рассматриваются подходы к комплексному типологическому описанию языков, которые представляют собой различные модели формализации представления их сопоставительно-типологических характеристик – в виде типологического паспорта языка, модели реферата. Указанные подходы свидетельствуют о стремлении к созданию систематизированных массивных баз данных языковой информации («Языки мира», «Всемирный атлас языковых структур»), которые могут служить основой для детализированного сопоставительно-типологического изучения языков, в том числе неблизкородственных и разноструктурных. Такие стандартизированные модели комплексного описания языков позволяют выявить культурную маркированность языковых единиц и целых языковых подсистем, а также являются актуальными с точки зрения их потенциала для автоматизации информационного поиска и использования методов компьютерной обработки лингво-типологических данных.

Ключевые слова: *типологический паспорт языка, культурно-типологический паспорт языка, модель реферата, сопоставительно-типологическое описание языков, детализация типологического паспорта, электронная база данных «Языки мира», Всемирный атлас языковых структур.*

Введение. В 1970–80-х гг. активно начали разрабатываться подходы и принципы комплексного типологического, а также сопоставительно-типологического описания языков мира. В результате появились различные модели экспликации содержательных характеристик языков с точки зрения формализации представления информации о них. Одной из самых первых таких моделей можно считать сопоставительное описание языков на основе их *типологических паспортов* [1; 2]. Идея создания такого паспорта принадлежит В.Д. Аракину, который применил в лингвистике практику, существующую в области техники, где широко применяются технические паспорта, которые представляют собой официальные документы, содержащие основные технические характеристики какого-либо механизма, агрегата и т.д., данные об их конструкции, технических свойствах, эксплуатационных качествах [3]. По аналогии с этим типологический паспорт языка представляет собой краткую и обозримую суммарную характеристику особенностей конкретного языка. Таким образом, типологический паспорт – это типовая схема, модель, отражающая в унифицированной форме сущностные характеристики языка.

Основная часть. Типологический паспорт языка, являясь основанием для сопоставительного исследования, может выполнять две взаимосвязанные роли. С одной стороны, типологический паспорт языка может стать результатом тщательного типологического изучения ведущих признаков структуры языка и выделения из них наиболее существенных, определяющих типологический облик данного языка [3]. На основе таких признаков и составляется типологический паспорт конкретного языка, который является конечной целью его описания в типологическом отношении. Примеры именно такого рода типологических паспортов мы и находим у В.Д. Аракина [3]. С другой стороны, типологический паспорт языка (как инвариантная модель с заранее заданным набором пунктов, отражающих наиболее существенные черты языковой структуры) может выступать в роли инструмента сопоставительно-типологического анализа языка, в ходе которого осуществляется своего рода «заполнение» пунктов паспорта. Иными словами, за отправную точку изучения языка в сопоставительно-типологическом плане берется некая абстрактная модель, составленная с учетом основных характеристик и особенностей, присущих человеческому языку вообще, что позволяет проследить (при его заполнении с применением данных конкретных языков), как проявляются эти характеристики и особенности в структуре конкретного языка. Отметим, что В.Д. Аракин разработал такой абстрактный типологический паспорт, в который необходимо включать следующие пункты:

- 1) состав фонемного инвентаря;
- 2) структура фонемного инвентаря;
- 3) ударение;
- 4) структура слога;

- 5) интонация;
- 6) состав грамматических категорий;
- 7) словосочетания;
- 8) предложение;
- 9) структура слова;
- 10) структура словарного состава;
- 11) словообразование;
- 12) общее заключение (типологическая характеристика языка, которая формулируется на основании учета всех типологических показателей).

Так как описание языка, представленное в типологическом паспорте, отражает лишь его основные типологические характеристики, то очевидно, что эти самые характеристики могут быть далее подвергнуты более детальному анализу с тем, чтобы выявить скрытые за поверхностной структурой языка его глубинные тенденции. Такая детализация типологического паспорта языка имеет несколько аспектов.

Можно пойти по пути количественного увеличения пунктов типологического паспорта, включив в него ряд дополнительных признаков, причем не только собственно языковых. Так, помимо основных фонетических, грамматических, синтаксических и лексических особенностей, типологический паспорт может содержать информацию об орфоэпической и орфографической нормах языка, его графическом своеобразии и т.д. В результате речь может идти о так называемом культурно-типологическом паспорте языка.

Очевидно, что такой подход к проблематике типологического паспорта позволяет получить представление лишь о внешнем облике языка, при этом его внутренняя организация остается достаточно скрытой и представлена не так наглядно. Однако если пойти по пути содержательного углубления уже имеющихся в нем пунктов, осуществить дальнейшую градацию содержания на уровне каждого пункта, то это даст возможность путем выделения внутри них подпунктов и аспектов рассмотрения языковых явлений выявить глубинные связи и закономерности структуры языка. Исследования В.А. Головина, который разработал алгоритм и критерии сопоставительного анализа лексической категории для типологического паспорта языка, иллюстрируют именно такой детализирующий подход [4]. Примером практического воплощения такого подхода могут служить исследования А.А. Игониной, которая, используя принципы и критерии, предложенные В.Д. Аракиным и В.А. Головиным, предприняла попытку более детально разработать критерии для типологического паспорта лексических категорий:

- I. Семантический критерий:
 - 1) значение, выражаемое лексической категорией;
 - 2) делимость/неделимость состава лексической категории на смысловые группы;
 - 3) степень зависимости от экстралингвистических факторов;
 - 4) открытость/замкнутость лексической категории.
- II. Формальный критерий:
 - 1) состав и тип маркеров: последняя основа сложного слова, аффиксы, полуаффиксы и т.д.;
 - 2) процентное соотношение способов образования лексических единиц;
 - 3) моноцентрическая/полицентрическая лексическая категория (под моноцентрической лексической категорией мы понимаем лексическую категорию, имеющую один тип маркеров, например, только суффиксы; под полицентрической лексической категорией мы подразумеваем лексическую категорию, имеющую различные типы маркеров, например суффиксы и последние основы сложных слов и т.д.);
 - 4) наличие/отсутствие способов выражения семантики лексической категории, выходящих за рамки уровня слова; при наличии последних их соотношение со способами выражения семантики лексической категории на уровне слова.

III. Функциональный критерий: типологически существенная лексическая категория в одном языке – слабая лексическая категория в другом языке [5, с. 101–102].

На основе практического применения данных критериев А.А. Игониной был разработан типологический паспорт лексической категории «место трудовой деятельности» в современном немецком языке [5]. Таким образом, типологический паспорт языка представляет собой действительно функционально эффективную модель стандартизированного сопоставительно-типологического описания языков.

Примерно в одно время с возникновением концепции типологического паспорта языка появилась другая модель комплексного описания языков, которая представлена разработкой многотомного энциклопедического издания «Языки мира» в Институте языкознания РАН. Поскольку «описание всех языков, независимо от их генетической принадлежности и социального статуса, выполняется в сопоставимой форме, по единым стандартам», то таким образом формируется база данных для любых сопоставительных и типологических исследований [6].

Статьи об индивидуальных языках написаны по единым стандартам в соответствии с типовыми схемами или «моделями реферата» [7]. По существу, типовые схемы представляют собой переработанный

и расширенный вариант типологического паспорта В.Д. Аракина, т.е. количественно и качественно «детализированный». Так, например, модель реферата содержит следующую информацию, не включенную в типологический паспорт:

- 1.1.0. Общие сведения
 - 1.1.1) варианты названия
 - 1.1.2) генеалогические сведения
 - 1.1.3) распространение; число говорящих
- 1.1.0. Лингвогеографические сведения
 - 1.2.1) общий диалектный состав
- 1.3.0. Социолингвистические сведения
 - 1.3.1) коммуникативно-функциональный статус и ранг языка
 - 1.3.2) степень стандартизации
 - 1.3.3) учебно-педагогический статус
- 1.4.0. Тип письменности
- 1.5.0. Краткая периодизация истории языка
- 1.6.0. Внутривидовые явления, обусловленные внешнеязыковыми контактами [6].

Даже беглый взгляд на приведенную типовую схему позволяет сделать вывод о том, что фактически в модели реферата реализована идея разработки культурно-типологического паспорта языка, о котором мы вели речь ранее в данной статье.

Кроме того, в модели реферата приводится более детальное внутривидовое описание фонетики, грамматики и лексики, что может быть проиллюстрировано сопоставлением следующих соотносительных пунктов типологического паспорта языка и модели реферата (см. полную схему: [8, с. 470–471]):

| | |
|--|---|
| <p>6. Состав грамматических категорий – данные по всем грамматическим категориям данного языка: категория падежа, рода, числа, детерминативности и др.; тип морфемных швов и характер морфем агглютинативных, флексий и др.</p> <p>10. Структура словарного состава – лексические категории, существующие в языке, общая характеристика отдельных лексико-грамматических разрядов, т.е. частей речи.</p> | <p>2.3.0. Семантико-грамматические сведения</p> <p>2.3.1. критерии выделения частей речи; категориальные способы выражения универсальных значений (общая характеристика)</p> <p>2.3.2. характер и способ выражения качественных именных классификаций</p> <p>2.3.3. категория числа и способы ее выражения</p> <p>2.3.4. падежные значения и их выражение; характер категорий принадлежности и их выражение</p> <p>2.3.5. характер и способ выражения качественных глагольных классификаций: залоговые, видовые и видовременные значения, наклонения, переходность, породы, версии</p> <p>2.3.6. дейктические категории и способы их выражения: категория лица и имени в глаголе, категория неопределенности/определенности в имени, категория времени в глаголе или в предложении, указание и пространственная ориентация, анафорические средства: выражение отрицания</p> <p>2.3.7. семантико-грамматические разряды слов</p> |
|--|---|

Таким образом, основными характеристиками как типологического паспорта, так и модели реферата являются стандартизованность и унифицированность представления информации о языках, вследствие чего обеспечивается межязыковая сопоставимость этой самой информации.

Практически одновременно с российским проектом «Языки мира» был создан «Всемирный атлас языковых структур» (The World Atlas of Language Structures – WALS) международным коллективом под руководством М. Хаспельмата, М. Драйера, Д. Гила и Б. Комри [9]. Атлас содержит данные по 2560 языкам (из всех семей) и 142 признакам, из которых 128 грамматические. Каждый признак может принимать от 2 до 9 значений, в среднем около 5. К недостаткам атласа относят тот факт, что не все языки описаны по отношению ко всем признакам, при этом отмечается, что в среднем язык в WALS имеет всего около 45 признаков [10, с. 22].

Анализ представленных выше моделей комплексного типологического описания языков свидетельствует о тенденциях к формализации представления информации с одной стороны, а также о стремлении к созданию систематизированных баз данных для сопоставительно-типологического изучения языков, с другой стороны. Такие базы данных могут оказаться особенно эффективными с точки зрения обеспечения необходимого инструментария:

- для упорядочения описания языков мира, выработке их классификаторов, составлению различных типологических и культурно-типологических атласов языков;
- для выделения типологически важных идиоэтнических/культурно-специфических и универсальных черт, а также выявления наличия лакун в структурах различных языков;

– для решения ряда методических задач, возникающих в процессе обучения родному и иностранному языкам (разработка методов и приемов изучения отдельных подсистем языка, в том числе посредством создания сопоставительных грамматик языков, контрастивных лексикографических материалов: контрастивных толково-переводных словарей, контрастивных словарей идиом, словарей безэквивалентной лексики и фразеологии);

– для решения вопросов, относящихся к сфере перевода и переводоведения, где часто из-за неадекватности языковых систем возникают проблемы, связанные с лакунарностью языковых средств и явлений и выбором способа ее компенсации.

Более того, описанные выше модели могут служить инструментом не только собственно лингвистического сопоставления. Так, на основе типологического паспорта или модели реферата можно проводить лингвогеографическое сопоставление различных языков или, в частности, сравнивать их учебно-педагогический статус. Кроме того, указанные выше базы данных – это своего рода информационные системы, которые, описывая язык в формализованных рамках, позволяют извлекать сведения лингвотипологического характера без непосредственного обращения к нормативным грамматикам, более громоздким лингвистическим справочникам и др.

Помимо прочего, стандартизированные модели описания языков представляются чрезвычайно актуальными с точки зрения реализации их потенциала для автоматизации информационного поиска, т.е. использования методов компьютерной обработки лингво-типологических данных. Такая работа уже осуществлена в рамках проектов «Языки мира» и «Всемирного атласа языковых структур», результатом реализации которых явились компьютерные базы данных, доступные в сети Интернет, и бумажные издания [6, 9].

Электронная база данных «Языки мира», созданная на основе текстовых материалов, содержит описания 315 языков Евразии по 3821 признаку, относящемуся к фонетике, морфологии и синтаксису [6; 10, с. 27]. Программная оболочка позволяет осуществлять поисковые, а также содержательные операции. Для проведения количественных типологических исследований важной является операция сравнения, позволяющая по двум выбранным языкам найти все признаки (и подсчитать их число), по которым эти языки совпадают и по которым они отличаются [10].

База данных Всемирного атласа (см. статистические данные выше) включает десятки тысяч справочных статей по различным признакам и языкам. Признаки классифицируются по тематике и по алфавиту, языки – по семьям и регионам. Поисковые средства позволяют находить нужную информацию по комбинации признаков. Кроме того, для каждого признака построена карта Земного шара, на которой кружочками разного цвета обозначены языки с различными значениями выбранного признака. Технически есть возможность генерировать новые карты с заданными свойствами [10, с. 25].

Иными словами, компьютерные базы данных, являясь удобным инструментом хранения и представления большого объема информации о языках, также обеспечивают возможность использования различных математических и статистических методов исследования для сопоставительно-типологического анализа большого массива лингвистических данных. При этом становится возможным трудоемкое сопоставление языков одновременно по десяткам или даже сотням параметров, а также разработка подходов к холистической классификации языков «на основе одновременного учета сотен и более грамматических свойств с применением объективных математических и компьютерных методов анализа» [11, с. 178].

Заключение. Таким образом, использование стандартизированных моделей комплексного описания языков обеспечивает сопоставимость языкового материала независимо от его происхождения, типологической классификации или структурного сходства/разнородности. Сопоставительно-типологическое изучение языков на основе таких типизированных моделей, оперируя заданным (широким) набором лингвистических характеристик, позволяет классифицировать их в рамках триады «универсалии – раритари¹ – уникалии». Обобщение таких данных даст возможность, во-первых, выявить те существенные характеристики, которые эксплицируют глобальные тенденции языкового развития, а во-вторых, систематизировав две последние категории, уточнить и сопоставить языковые картины мира на базе широкого диапазона языков, в том числе неблизкородственных и разноструктурных. Кроме того, использование созданных на основе описанных выше моделей компьютерных баз данных позволяет расширить диапазон подходов к классификации языков с учетом результатов анализа большого количества параметров.

ЛИТЕРАТУРА

1. Бетенья, Е.П. Сопоставительное изучение языков на основе типологического паспорта / Е.П. Бетенья; Минский гос. лингвист. ун-т. – Минск : МГЛУ, 1997. – 38 с.
2. Бетенья, Е.П. От типологического паспорта языка к модели реферата // *Личность – слово – социум : материалы 4-й Междунар. науч.-практич. конф.*, Минск, 14–16 апреля 2004 г. / Отв. ред. В.В. Фалалеев. – Минск : ЧУП «Паркус плюс», 2004. – С. 69–72.

¹ Термин предложен Е.И. Ярославцевой [12].

3. Аракин, В.Д. Типология языков и проблема методического прогнозирования / В.Д. Аракин. – М. : Высш. шк., 1989. – 158 с.
4. Головин, В.А. Теория типологических паспортов В.Д. Аракина и актуальные проблемы лингвистической паспортизации / В.А. Головин // Вопросы лингвистической типологии, истории языка и лингводидактики. – М. : Владос, 1994. – С. 5–16.
5. Игонина, А.А. Типологический паспорт лексической категории / А.А. Игонина // Известия Росс. гос. пед. ун-та им. А.И. Герцена. – 2007. – Т. 16, № 40. – С. 100–103.
6. Компьютерная база данных «Языки мира» [Электронный ресурс]. – Режим доступа: <http://www.lingvarium.org/langworld/>. – Дата доступа: 22.06.2020.
7. Виноградов, В.А. База данных «Языки мира» как инструмент лингвистического исследования / В.А. Виноградов, А.И. Новиков, Е.И. Ярославцева // Вопросы языкознания. – 2003. – № 3. – С. 3–14.
8. Языки мира. Германские языки. Кельтские языки / глав. редкол.: В.Н. Ярцева, В.М. Солнцев, Н.И. Толстой. – М. : Academia, 2000. – 472 с.
9. The World Atlas of Language Structures Online [Electronic Resource]. – Mode of access: <https://wals.info>. – Date of access: 22.06.2020.
10. Поляков, В.Н. Компьютерные модели и методы в типологии и компаративистике. / В.Н. Поляков, В.Д. Соловьев. – Казань : Казан. гос. ун-т им. В.И. Ульянова-Ленина, 2006. – 208 с.
11. Поляков, В.Н. Типологическая схожесть языков как метод изучения языковой эволюции / В.Н. Поляков // Вестник РГГУ. Сер. Филол. науки. Языкознание. Вопросы языкового родства. – 2010. – № 16 (59). – С. 177–198.
12. Ярославцева, Е.И. Компьютерная база данных «Языки мира» и ее возможные применения : автореф. дисс. ... д-ра филол. наук : 10.02.21 / Е.И. Ярославцева; Ин-т языкознания РАН. – М., 2005. – 40 с.

Поступила 23.06.2020

TOWARDS THE HISTORY OF FORMALIZING OF THE COMPARATIVE AND TYPOLOGICAL LANGUAGES DESCRIPTION

E. BETENYA

The article deals with approaches to complex typological description of languages which represent different models of formalizing the presentation of their comparative and typological characteristics of different languages in the form of a typological language passport, abstract model. These approaches illustrate the motivation to create systematized massive databases of language information (World Languages, The World Atlas of Language Structures), which may serve as the basis for a detailed comparative and typological study of languages, including genetically non-related and structurally heterogeneous ones. Such standardized models for a comprehensive description of languages make it possible to reveal the culture-bound aspect of linguistic units and entire linguistic subsystems. In addition, they present relevance as regards their potential for automation of information retrieval and use of computer processing methods for linguistic and typological data.

Keywords: *typological language passport, cultural and typological language passport nacpnm, models of formalizing, comparative and typological characteristics, electronic database World Languages, The World Atlas of Language Structures.*