

Г. М. КЕРТ, В. А. ЛЕБЕДЕВ (Петрозаводск)

ОБ ИССЛЕДОВАНИИ ТОПОНИМИИ СЕВЕРА ЕВРОПЕЙСКОЙ ЧАСТИ СССР

Топонимия Севера Европейской части СССР как по структурным признакам, так и по семантике представляет собой в известной степени результат исторических процессов, протекавших на этой территории. Согласно археологическим данным, здесь проходили длительные миграции и это не могло не оставить следа в топонимии. Четко прослеживается слой топонимов, которые невозможно этимологизировать на базе существующих в настоящее время языков уральской, тюркской или индоевропейской семьи. Наиболее мощным пластом топонимии является уральский, представленный прибалтийско-финскими, саамским, коми, угорскими и самодийскими языками. Начиная с XI в., а может быть и раньше, топонимия Севера Европейской части СССР подвергается все большему русскому влиянию.

1. Задачи исследования

1. Выявление субстратного слоя топонимии, его фонетических и морфологических особенностей, ареала распространения. Установление степени однородности, повторяемости топонимов, их плотности на условной единице территории.
2. Определение ареалов топонимии, этимологизируемой с помощью живых языков, распространенных на территории Европейского Севера СССР, структурных типов топонимов и их семантики: 1) прибалтийско-финская, 2) саамская, 3) коми, 4) угорская, 5) самодийская, 6) русская (славянская) топонимия.
3. Выявление ареалов смешения топонимии, зон топонимии двух и более языков, соотношения частот различных топонимов в разных ареалах.
4. Определение абсолютных и относительных частот топонимов каждого языка на единицу площади.
5. Распространение структурных типов топонимов по различным ареалам: 1) однокорневые, 2) двухкорневые, 3) состоящие из трех и более корней, 4) суффиксальные, 5) прочие.
6. Сопоставление топонимических данных с данными истории (акты, летописи, писцовые книги), археологии, этнографии, фольклора.
7. Выявление общих принципов адаптации субстратных топонимов.
8. Выявление изоглосс однотипных явлений как в семантическом, так и формальном отношении.
9. Выявление детерминативов различных языков.
10. Выявление топонимических типов (по А. Матвееву) и общих топооснов.
11. Исследование соотносительности семантики топонима с характеристикой географического объекта — совпадение или расхождение.
12. Сопоставление субстратных топонимов и апеллятивных заимствований.
13. Наличие калек в субстратной топонимии, различных фонетических вариантов, гибридных сочетаний.

14. Исследование топонимии по семантическим классам (классификация по системе понятий Рудольфа Халлига и Вальтера фон Вартбурга).
 15. Выяснение системных связей (семантических оппозиций) в топонимии.
 16. Выявление соотношения смоделированного словаря топонимов (по методике А. Матвеева) с сеткой субстратных топонимов.
 17. Выявление топонимов, отражающих тектонические движения земной коры, катастрофические явления природы, внезапные исчезновения озер, изменение направления течения рек и др.
- Разумеется, могут быть поставлены и другие задачи.

2. Принципы формализации топонимов

Значительная часть названных задач может решаться с помощью формализованных методов, что в свою очередь позволяет поставить вопрос о применении ЭВМ в исследованиях топонимии. Целесообразность этого обуславливается прежде всего большой трудоемкостью подсчета числовых характеристик топонимии, разнообразием вариантов задач, а также монотонностью работы, требующей большого внимания от исполнителей. Применение ЭВМ высоко эффективно для получения различных тематических выборок, сопоставления значений топонимов и апеллятивов по семантическим признакам, редактирования и во многих других случаях. Немаловажным обстоятельством является и то, что разработанное программное обеспечение может многократно применяться для исследования топонимии больших территорий, различных языков и при постепенном накоплении топонимической лексики.

Каждый топоним обладает суммой твердо заданных признаков, которые можно формализовать. К ним относятся географические, графические (перевод звукового или графического кода в единый графический), структурные и семантические. Географический признак — это местонахождение топонима на местности. Формализация графического облика топонима также не составит больших трудностей. Транскрипция будет максимально упрощена, чтобы в ЭВМ можно было заложить данные языков различных звуковых типов. Более того, большинство топонимов будет поступать в адаптированном русском языке виде. Сложнее дело обстоит с формализацией структурных типов, поскольку в орбиту исследования попадут топонимы разносистемных языков. Учитывая, что структура топонимов составляет малую часть структуры языка, надо полагать, что и эта задача разрешима. Наибольшие трудности представит формализация семантических типов, однако и она поддается решению.

Какие же задачи и каким образом целесообразно решать при помощи ЭВМ? Прежде всего это подсчет абсолютных и относительных частот появления топонимов по различным участкам территории, их сравнение и сопоставление (например, выявление изоглоссе); построение ареалов топонимии с наложением на карту; анализ топонимов по семантическим классам лексики, по способам образования топонимов с подсчетом числовых характеристик и т. д.

Ввиду многообразия задач и их вариативности по условиям и исходным данным, а также для обеспечения возможности наращивания исследуемой топонимии, перехода от одних территорий к другим и исследования многоязычной топонимии наиболее целесообразным представляется построение базы данных по топонимии с последующим решением на ее основе разнообразных задач. Преимущества создания баз данных заключаются в следующем: во-первых, данные в память ЭВМ вносятся однократно и, во-вторых, они могут многократно использоваться для решения всевозможных задач. Это дает значительное сокращение затрат труда на выполнение всех работ. К недостаткам можно отнести необходимость тщательного проектирования структуры и содержания и относительную длительность создания баз данных. Однако в этом же заключается и достоинство, так как требуется тщательная проработка методики исследований.

Рассмотрим предварительные структуры и содержание базы данных топонимии, предварительные потому, что в процессе рабочего проектирования предлагаемая структура и содержание могут быть изменены.

Сначала перечислим основные требования к базе данных. Она должна обеспечивать возможности структурного и семантического анализа топонимов, их отнесения к соответствующей единице площади, топографическому объекту и классу, указывать языковое происхождение топонима, способы образования, грамматические признаки. База данных должна содержать словарь лексических основ (корней), аффиксов и падежных окончаний с указанием значений и принадлежности к определенному языку по написанию и происхождению. Данные могут быть в латинском или русском алфавите и в числовой форме, при этом для аффрикат и разделения лексических составляющих следует подобрать специальные знаки (поскольку надстрочные и подстрочные знаки для ЭВМ недопустимы, некоторые буквы (фонемы) будут выражены двумя символами). Возможны и дополнительные требования, которые в настоящее время не могут быть сформулированы.

Итак, предлагается структура базы данных топонимии, которая содержит таблицы-словари: основ, аффиксов, окончаний, собственно топонимов и покрытия топонимами карты с разделением их по языкам. Например, будут представлены таблицы топонимов на русском, саамском и прибалтийско-финских языках. Разумеется, база данных может наращиваться включением таблиц топонимов на других языках (например, коми, ненецкий, мансийский). Во время решения различных задач таблицы могут быть связаны между собой динамически соответственно различным условиям. Строки таблиц, составляющие их содержание, могут пополняться и изменяться в процессе эксплуатации базы данных. Для этого предусматриваются специальные операции в программном обеспечении. Возможно включение в таблицы дополнительных граф, значения которых (данные) вносятся извне или вычисляются на основе уже имеющихся.

Структура словаря топонимов

Язык представления	Номер по порядку	Топоним	Язык происхождения	Способ образования	Русский эквивалент (калька)	Число лексических компонентов	Соответствие объекту	Пункт сбора материала или источник	Апеллятив
--------------------	------------------	---------	--------------------	--------------------	-----------------------------	-------------------------------	----------------------	------------------------------------	-----------

Значение графы «Номер по порядку» будет автоматически вычисляться при вводе данных. Таблицу целесообразно сохранять упорядоченной 1) лексикографически по топонимам и 2) по порядковым номерам.

Покрытие карты топонимами

Язык представления	Элемент карты	Квадрат	Административная область	Условное обозначение класса объектов	Номер по порядку
--------------------	---------------	---------	--------------------------	--------------------------------------	------------------

Главная особенность этой таблицы — способ представления площади карты с разбиением ее на квадраты (что принято в топографии), а их в свою очередь на элементарные квадраты (элементы). Величина элемента определяется делением строки квадрата на 10 или 100 частей, что будет установлено позднее. Элемент содержит один топоним, несколько или ни одного. Соответственно таблица располагает одной или несколькими строками на элемент. В случае отсутствия топонима строка все равно записывается, но в двух последних графах ставятся прочерки. По содержанию этой таблицы будут определяться ареалы топонимов, основ, способов образований и т. д. и распечатываться в виде схематической карты. Формирование ее составит определенные трудности, но они вполне преодолимы.

Для обеспечения структурного и семантического анализа топонимов, а также расчета различных числовых характеристик по их компонентам предназначены таблицы основ (корней), аффиксов и окончаний с указанием необходимых характеристик.

Структура словаря основ

Язык представления	Основа	Русский эквивалент (калька)	Семантический класс	Язык происхождения	Часть речи
--------------------	--------	-----------------------------	---------------------	--------------------	------------

Структура словаря аффиксов

Язык представления	Аффикс	Тип	Значение
--------------------	--------	-----	----------

На основе представленной структуры данных будет разрабатываться программное обеспечение для ее создания и введения, а также последующего решения задач. Предполагается использование ЭВМ типа ЕС как наиболее распространенной в настоящее время. В качестве программной основы будет принята какая-нибудь из готовых систем управления базами данных. Программное обеспечение должно позволять работать с базой данных лингвистам в интерактивном режиме при помощи дисплея.