

Горячева Мария Александровна, Горячев Николай Сергеевич

ИНФОРМАЦИОННАЯ СИСТЕМА "ЯЗЫКОВОЕ РАЗНООБРАЗИЕ РОССИИ": ОСНОВНЫЕ ИТОГИ ПРОЕКТА

Статья посвящена проекту Научно-исследовательского центра по национально-языковым отношениям Института языкознания РАН Информационная система "Языковое разнообразие России". В нем представлены данные о современном состоянии и динамике социальных функций около 80 языков Российской Федерации. Система вводит в научный оборот большой объем четко структурированной информации, представленной на основе принципа удобства и открытости использования. Она дает возможность быстрого обновления данных, расширения состава языков, а также позволяет использовать ее как основу для создания подобных проектов как в отдельных регионах страны, так и в других странах.

Адрес статьи: www.gramota.net/materials/2/2016/11-1/31.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2016. № 11(65): в 3-х ч. Ч. 1. С. 110-112. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2016/11-1/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net
Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: phil@gramota.net

УДК 81.272

Статья посвящена проекту Научно-исследовательского центра по национально-языковым отношениям Института языкознания РАН Информационная система «Языковое разнообразие России». В нем представлены данные о современном состоянии и динамике социальных функций около 80 языков Российской Федерации. Система вводит в научный оборот большой объем четко структурированной информации, представленной на основе принципа удобства и открытости использования. Она дает возможность быстрого обновления данных, расширения состава языков, а также позволяет использовать ее как основу для создания подобных проектов как в отдельных регионах страны, так и в других странах.

Ключевые слова и фразы: информационная система языков; социолингвистика; языковая ситуация в России; демографическая мощьность языков; коммуникативная мощьность языков; коммуникативные сферы; функциональное развитие языков; языковая ситуация; функциональный статус языков.

Горячева Мария Александровна, к. филол. н.
Институт языкознания Российской академии наук, г. Москва
masha22275@list.ru

Горячев Николай Сергеевич, к. физ-мат. н.
Институт проблем химической физики Российской академии наук, г. Черноголовка
gor_ns@icp.ac.ru

ИНФОРМАЦИОННАЯ СИСТЕМА «ЯЗЫКОВОЕ РАЗНООБРАЗИЕ РОССИИ»: ОСНОВНЫЕ ИТОГИ ПРОЕКТА

*Статья подготовлена при поддержке гранта РГНФ № 14-04-12023
Информационная система «Языковое разнообразие России» (рук. М. А. Горячева).*

В настоящее время Интернет предоставляет широкие возможности для научного сообщества. Прежде всего, это быстрое нахождение научной информации, поиск ученых и научных групп, занимающихся сходной проблематикой, а также использование интернет-ресурсов как площадки для научного общения. Эти возможности с каждым годом расширяются: все больше журналов публикуют статьи своих авторов в Интернете, оцифровываются библиотеки и архивы, создаются тематические сайты, интернет-публикация приравнивается к печатной. Нетрудно предположить бурный скачок во всех областях научного знания за счет кумулятивного эффекта.

Особую роль в процессе расширения информационных возможностей научного сообщества играют информационные системы (ИС). ИС – система, предназначенная для хранения, поиска и обработки информации и для своевременного удовлетворения конкретных информационных потребностей в рамках определенной области знаний. Первыми ее начали осваивать представители технических и естественных наук, но в дальнейшем стали успешно использовать и гуманитарии. Как пример больших информационных массивов можно привести данные по солнечной активности; информацию, поступающую с телескопов; данные по Большому адронному коллайдеру. В области гуманитарных наук, в частности языкознания, ИС также становятся все более распространенными. Так, в Институте языкознания в этом году закончился Лингвистический проект «Глобальная лексикостатистическая база данных» (ГЛБД). Это проект Московской школы компаративистики, осуществляемый при поддержке программы Института Санта-Фе «Эволюция человеческого языка» (Evolution of Human Language, EHL) и проекта «Вавилонская башня», начатого в 1998 году С. А. Старостиным. Цель проекта ГЛБД – составление стандартизованных списков слов базисной лексики возможно большего числа языков мира и их диалектов, а также списков реконструированных праформ многочисленных праязыков, и обеспечение широкого публичного доступа к этим спискам [1]. В западном языкознании есть ряд проектов, задачей которых является описание грамматических признаков языков по единой классификации. Можно назвать «Этнолог» (Ethnologue: Languages of the World) и недавний проект «Реестр Лингвосферы» (The Linguasphere Register of the World's Languages and Speech Communities). Крупным международным исследовательским проектом в области лингвистической географии является Лингвистический атлас Европы (ЛАЕ).

Следует отметить значение такого свойства ИС, как возможность сделать содержащуюся в них информацию открытой для широкого круга пользователей. Особенно это важно в гуманитарной сфере, в частности, в таких науках, как история, социология, политология, социальная лингвистика и т.п. Доступная и объективная информация поможет предотвратить возникновение и развитие многих конфликтов, искажение исторических фактов, манипулирование общественным сознанием.

С другой стороны, необходимым условием воссоздания наиболее приближенной к реальной действительности картины является высокий уровень достоверности представляемых фактов. К сожалению, ряд располагающихся в сети энциклопедий, специализированных сайтов и т.п. не может гарантировать соблюдения этого условия. Примером тому может быть популярная и всеми любимая Википедия, которая «пишется народом». Поэтому наряду с открытостью для использования информации необходимо соблюдение условия ее достоверности. Это достигается, с одной стороны, подбором коллектива, являющегося профессионалами в данной области, и, с другой стороны, отсутствием возможности ее изменения и дополнения пользователями напрямую, минуя ее авторов. Возможно, отсутствие этой возможности в системе является устаревшим подходом, но это поможет защитить ее от неверной информации. Обратная связь в таком случае осуществляется через электронную почту.

Авторы ИС «Языковое разнообразие России» – сотрудники Научно-исследовательского центра по национально-языковым отношениям (НИЦ НЯО). Он создан в 1989 г. по решению Президиума РАН на базе сектора социолингвистики Института языкознания РАН.

Новый проект Научно-исследовательского центра по национально-языковым отношениям Института языкознания РАН «Языковое разнообразие России» относится к социолингвистическим обобщающим информационным трудам. В нем представлены данные о современном состоянии и динамике социальных функций около 100 языков коренных народов Российской Федерации и о языковой ситуации в регионах страны.

Система направлена на создание возможности быстрого доступа к большому массиву статистических и фактических данных по функционированию языков России, с возможностью их последующего обновления и дополнения. Это позволит исследованию по выявлению и теоретическому осмыслению тенденций функционального развития этих языков в период 1989-2016 гг. Станет возможным проследить действие в России основных тенденций: расширение или утрата ряда социальных функций или даже уход из употребления некоторых языков народов России. Полученная информация обогатит теоретическую базу социолингвистики и может стать основой для дальнейшего изучения языковых процессов, разработки принципов языковой политики. Таким образом, информационная система позволит не только ввести в научный оборот большой массив четко структурированной информации, поданной на основе принципа удобства и открытости использования, но и создаст предпосылки для дальнейшего теоретического развития социолингвистической науки, особенно в проведении исследований, требующих обработки большого объема информации. В частности, проект внесет ощутимый вклад в развитие социолингвистической систематики языков.

Поскольку Информационную систему можно назвать проектом-спутником к другим проектам НИЦ НЯО, мы сочли логичным максимально приблизить его структуру к следующим работам центра: к международному российско-канадскому труду «Письменные языки мира. Языки Российской Федерации» в двух томах и соответствующему разделу Социолингвистической энциклопедии «Язык и общество» (в печ.). Коллективный труд НИЦ НЯО ИЯз РАН «Письменные языки мира. Языки Российской Федерации» был выполнен в рамках международного проекта «Письменные языки мира: степень и способы использования» (WLW 1988-1998) с привлечением ученых из республик и регионов России [4; 5]. Тогда же была разработана «Карта описания языка», заложившая основу принципов рубрикации и в данном проекте. Как особый прием изучения, в ней использовался социолингвистический портрет языка [3].

Итак, информационная система находится в русле традиций трудов центра, а также выходит в интернет-пространство в удобной для исследователей форме [2]. Такой подход дал возможность обеспечить преемственность трудов центра. При этом был внесен ряд изменений и дополнений, что объясняется, с одной стороны, задачами максимально оптимизировать структуру с целью удобства пользования и, с другой стороны, детализировать ее с целью предоставить более подробную информацию. Так, ряд доминантных коммуникативных сфер представлен более подробно, что даст возможность увидеть динамику уровня использования в них языков. Нашла отражение и информация по языкам нового региона Российской Федерации – Крыма (русский, украинский и крымско-татарский языки).

В социолингвистике можно констатировать нехватку подобных информационных систем, что в значительной мере тормозит как теоретическое развитие науки (особенно проблемы ее систематики, функциональной типологии языков и мн. др.), так и практическое применение ее достижений в других областях (например, в политологии, конфликтологии, педагогике, изучающей проблемы регионального образования, и т.д.).

Как и любой подобный проект, создание ИС имеет следующие этапы: разработка концепции (в частности, формулировка идеи, формирование команды и др.), подготовка технического предложения, проектирование, разработка, ввод в эксплуатацию. Мы сейчас находимся на этапе заполнения системы статистической и фактической информацией.

Информационная система будет представлена в свободном доступе на сайте Института языкознания РАН, поэтому сфера использования проекта не ограничена научной аудиторией: он рассчитан на широкий круг пользователей Интернета, интересующийся языковой ситуацией в России. Прежде всего, она будет востребована лингвистами, а также политологами и педагогами, занимающимися региональными проблемами образования и культуры.

Пользователь, на которого ориентируются создатели ИС, не обязан обладать какой-либо квалификацией в области вычислительной техники. Поэтому интерфейс, предлагаемый ему, должен быть простым и удобным, с легко осваиваемыми функциями, не перегруженным лишними действиями. Исходя из данного условия, нам нужно было решить две основные задачи: разработку базы для хранения информации и удобного для пользователей интерфейса. ИС создана на основе свободной реляционной системы управления базами данных MySQL, которая была выбрана потому, что является оптимальным решением для малых и средних проектов, входит в состав многих серверов и обладает значительной гибкостью за счет большого количества поддерживаемых типов таблиц. Результаты будут представлены в виде статистических запросов с настраиваемыми параметрами и возможностью вывода на печать. Внешнее оформление сайта было построено на основе системы управления содержимым Drupal. Эта система основана на свободном программном продукте и активно развивается.

Представлены следующие информационные блоки:

- общая характеристика языка;
- общие сведения о носителях данного языка;
- сведения о владении языками;
- письменность и литературный язык;
- социальные функции языка;

- функциональные характеристики языков;
- заключение.

Кратко охарактеризуем содержание каждого пункта.

В пункт 1 «Общая характеристика языка» входят следующие подпункты: названия языка (официальное, самоназвание, устаревшие); юридический статус языка (с указанием документов); место в генетической классификации; диалектное членение; типологическая характеристика; внутренняя структура языка (краткий очерк грамматической структуры и лексики).

Пункт 2 «Общие сведения о носителях языка» включает сведения об этнониме, языковой общности, традиционной территории проживания, численности этноса (в том числе за ее пределами), численности проживающих в городской и сельской местности.

Пункт 3 «Сведения о владении языками» представляет статистические данные переписей населения 1989, 2002 и 2010 гг. о владении представителями данного этноса родным и русским языками.

Пункт 4 «Письменность и литературный язык» описывает историю создания письменности (первые тексты, тип языка: старописьменный/младописьменный/с прерванной традицией, смена графической основы), алфавит, дает краткую характеристику литературного языка, выделяет периодизацию его становления, а также перечисляет представленные в языке основные стили и жанры.

В пункте 5 «Социальные функции языка» приводится большой объем статистической и фактической информации по использованию языка в основных регулируемых и нерегулируемых государством сферах общения. Это сферы: образования (дошкольное, школьное, среднее профессиональное и высшее, а также использование языка в учебной и методической литературе); науки; средств массовой информации (газеты и журналы, радио и телевидение); художественной литературы; театра и кино; законодательства; административной деятельности; религии (использование в обрядах, переводы священных текстов); судопроизводства; сельскохозяйственного, промышленного производства; обслуживания и торговли; транспорта и Интернет. Сведения о применении языка в нерегулируемых сферах общения (семейное, повседневное бытовое и межличностное) включают результаты опросов, анкетирований и личных наблюдений исследователей. В Заключении приводятся краткие выводы.

Проект имеет широкие возможности и перспективы развития. Прежде всего, это возможность быстрого обновления данных. Это обязательное условие для полноценного функционирования информационных систем. Постоянно появляются новые статистические и фактические материалы. Перспективным направлением развития проекта является и расширение состава языков: включение языков этнических групп и диаспор. Также данную ИС можно использовать как основу для создания подобных проектов по социолингвистическому описанию языков как в отдельных регионах страны, так и в других странах.

Список литературы

1. **Глобальная лексикостатистическая база данных: общее описание** [Электронный ресурс]. URL: <http://starling.ginnet.ru/new100/GLDr.htm> (дата обращения: 27.09.2016).
2. **Горячева М. А.** Информационная система «Языковое разнообразие России» – проект-спутник к трудам Научно-исследовательского центра по национально-языковым отношениям // *Языковая политика и языковые конфликты в современном мире: доклады и сообщения* (16-19 сентября 2014 г.). М.: Тезаурус – Языки народов мира, 2014. С. 309-314.
3. **Михальченко В. Ю.** Социолингвистический портрет письменных языков России: методы и принципы исследования // *Методы социолингвистических исследований*. М.: МГПИ «Доминант», 1995. С. 205-213.
4. **Письменные языки мира. Языки Российской Федерации. Социолингвистическая энциклопедия:** в 2-х кн. / отв. ред. В. Ю. Михальченко. М.: Academia, 2000. Кн. 1. 656 с.
5. **Письменные языки мира. Языки Российской Федерации. Социолингвистическая энциклопедия:** в 2-х кн. / отв. ред. В. Ю. Михальченко. М.: Academia, 2003. Кн. 2. 848 с.

THE INFORMATION SYSTEM “LANGUAGE DIVERSITY OF RUSSIA”: THE MAJOR RESULTS OF THE PROJECT

Goryacheva Mariya Aleksandrovna, Ph. D. in Philology
Institute of Linguistics of the Russian Academy of Sciences, Moscow
masha22275@list.ru

Goryachev Nikolai Sergeevich, Ph. D. in Physical and Mathematical Science
Institute of Problems of Chemical Physics of the Russian Academy of Sciences, Chernogolovka
gor_ns@icp.ac.ru

The article is devoted to the project of the Research Center on Ethnic and Language Relations of the Institute of Linguistics, RAS **Information System “Language Diversity of Russia”**. It presents the data on the current state and dynamics of the social functions of about 80 languages of the Russian Federation. The system introduces into scientific circulation a large amount of well-structured information provided on the basis of the principle of ease and openness of use. It allows fast data updating, the enlargement of languages structure, and also gives the possibility of using it as the basis for the creation of similar projects both in individual regions of the country and in other countries.

Key words and phrases: information system of languages; sociolinguistics; language situation in Russia; demographic power of languages; communicative power of languages; communicative spheres; functional development of languages; language situation; functional status of languages.