

Лингвогеографическая информационная система «Диалект» и
лингвистические исследования

Е. А. Жданова, И. С. Соломенников

Ижевский государственный технический университет, Удмуртский
государственный университет, Ижевск, Россия

Как отмечалось ранее [Жданова 2006: 47-50; Соломенников 2006: 161-163], лингвогеографическая информационная система «Диалект» (ЛГИС «Диалект») предоставляет возможность:

- 1) вводить и редактировать лексический материал, собранный по вопросам Лексического атласа русских народных говоров (ЛАРНГ);
- 2) осуществлять выборку данных, готовить ее для картографирования, корректируя состав картографируемых объектов в зависимости от интересов и задач пользователя;
- 3) визуализировать данные в виде лингвогеографической карты, на которой с помощью знаков могут быть представлены сложные диалектные соответствия.

В настоящее время ЛГИС «Диалект» значительно доработана. В дополнение к лексическому модулю в системе появился фонетический модуль, вопросник которого основан на программе Диалектологического атласа русского языка. Разработан специальный фонетический шрифт, а также интерфейсы для ввода транскрипции слов. С технической точки зрения фонетический модуль почти идентичен диалектному, но в нем ввиду специфики материала хранятся меньше свойств и значений данных. Также переработан модуль выборки и визуализации информации на карте – теперь он позволяет строить фонетические карты.

В отличие от предыдущей версии, которая позволяла производить выборку только по одному критерию – одному из вопросов ЛАРНГ, сейчас имеется возможность осуществлять отбор данных одновременно с учетом нескольких параметров:

- 1) сведения об информаторе, от которого было записано слово, что дает возможность выбрать одного или нескольких информаторов,

удовлетворяющих одному или нескольким значениям параметров «Пол», «Профессия», «Образование», «Национальность», «Дата рождения» и др.;

2) сведения о населенном пункте, где происходил сбор материала, что позволяет скорректировать выборку в зависимости от количества жителей в том или ином месте сбора материала, от наличия/отсутствия населения другой национальности и от некоторых других, существенных для диалектной выборки параметров; кроме того, пользователю предоставляется возможность увидеть, по каким разделам Программы ЛАРНГ был проведен опрос в выбранном населенном пункте;

3) сведения о собирателе, зафиксировавшем ту или иную единицу, что дает возможность отобрать данные, записанные конкретными людьми, и, если необходимо, провести анализ с учетом степени достоверности материала.

Чрезвычайно важной представляется новая возможность осуществлять выборку по маске слова, что дает возможность получить сведения об употребительности/неупотребительности данной номинации на картографируемой территории и о том, для обозначения чего она используется.

Так, например, введя в поле для поиска маску *хороший*, по базе данных русских говоров Удмуртии мы получаем список из 16 обозначений аналогичного вида, являющихся ответами на разные вопросы Программы ЛАРНГ. Первичный анализ данных может быть проведен с помощью дополнительной информации, такой, как (а) код соответствующего вопроса Программы, (б) имена и фамилии информатора и собирателя, (в) населенный пункт, (г) формулировка и тип вопроса, ответом на который является слово *хороший*, (д) его грамматическая характеристика (имя существительное/прилагательное), (е) сведения об информаторе (пол, образование, национальность) и населенном пункте (район, область). Такая информация либо отображается в явном виде в списке результатов выборки, либо вынесена во всплывающие подсказки, активирующиеся при наведении курсора мыши на слово.

Новинкой, доступной только для зарегистрированных пользователей, является функция сохранения созданных выборок для последующего использования. Данная функция сделана для удобства исследователей и позволяет неоднократно возвращаться к уже проанализированным данным для их повторного анализа, сравнения.

Следующим нововведением является замена собственных обозначений на те, что используются в Институте лингвистических исследований РАН. Были созданы наборы более чем из 300 значков в нескольких цветовых вариантах.

Использование ЛГИС «Диалект» позволяет получить такие данные, которые могут быть сразу интерпретированы. Так, выборка слова *хороший* показала, что в русских говорах Удмуртии оно значит 'съедобный' (Какмож, Первомайский, Девятово, Завьялово), 'вкусный' (Первомайский), 'ясный, солнечный' (Гура), 'умный человек' (Перевозное), 'добрый' (Перевозное), 'застенчивый' (Перевозное), 'делающий добро другим, добрый человек' (Камское), 'лен высокого качества' (Верхняя Позимь), 'вежливый, обходительный' (Кез). Сравнение этого материала с данными, полученными по маске *плохой*, давшими, помимо прочих реализаций, номинацию *плохой* в значении 'несъедобный' (Какмож, Девятово, Завьялово) почти во всех тех же населенных пунктах, где отмечено *хороший* в значении 'съедобный'. Это свидетельствует об определенной системности развития семантики слов в говорах исследуемой территории, что дает основание говорить о наличии определенных словообразовательных моделей семантической деривации в русских говорах Удмуртии.

Подобные выборки могут применяться для изучения омонимии в диалектном языке и, в частности, для исследования проблемы семантического словообразования в русских говорах Удмуртии.

Перспективами дальнейшего развития системы являются:

- интеграция ее с системой «Google Maps», преимущества которой в том, что она предоставляет хорошие возможности по отображению местности в нескольких вариантах, имеет гибкий механизм

масштабирования, а также в том, что позволяет отображать на карте всю территорию планеты;

- поддержка многоязычности в интерфейсе системы, что позволит работать с ней не только русскоязычным лингвистам, но и ученым из других стран мира;
- дальнейшее совершенствование модуля запросов – внедрение в него новой функциональности, что позволило бы более гибко и полно производить анализ собранного материала.

Список литературы

Соломенников 2006 – *Соломенников, И. С.* Лексический атлас русских говоров Удмуртии в Интернет / *И. С. Соломенников* // Современные информационные технологии и письменное наследие: от древних рукописей к электронным текстам : материалы междунар. науч. конф., Ижевск, 13-17 июля 2006г. / Отв. ред. В. А. Баранов. – Ижевск : Изд-во ИжГТУ, 2006. – С. 161-163.

Жданова 2006 – *Жданова, Е. А.* Применение лингвогеографической информационной системы «Диалект» в лингвистических исследованиях / *Е. А. Жданова* // Современные информационные технологии и письменное наследие: от древних рукописей к электронным текстам : материалы междунар. науч. конф., Ижевск, 13-17 июля 2006г. / Отв. ред. В. А. Баранов. – Ижевск: Изд-во ИжГТУ, 2006. – С. 47-50.

Dialect, a linguistic-geographical information system, and linguistic research

Ekaterina A. Zhdanova, Igor' S. Solomennikov

Izhevsk State Technical University, Udmurtia State University, Izhevsk, Russia

This paper discusses Dialect, a linguistic-geographical information system developed at Udmurt State University in 2005. The system continues to be improved, expanding researchers' opportunities in the field of studying dialectical speech and language.