Характеристика текстов

В данном разделе представлен Устный корпус осетинского языка (УКО), в который входят тексты на иронском, дигорском и уаллакомском диалектах. Тексты представляют собой свободные диалоги и рассказы жителей разных сёл Северной Осетии. Среди иронских текстов присутствует также некоторое количество аудиозаписей радиопрограммы "Одноклассники"
(
радиостанции "Алания").
Все тексты имеют поморфемный разбор, перевод и грамматические комментарии на русском и английском языках. Тексты также снабжены аудиозаписями. Буквальный перевод, приведённый для предложений в устных текстах, намеренно сделан близким к осетинскому оригиналу, и поэтому может выглядеть несколько искусственным.
Общий объем Устного корпуса осетинского языка - 77000 словоупотреблений, из них на иронском диалекте - 60000, на дигорском диалекте - 13000, на уаллакомском диалекте - 4000.

Участники проекта

Запись текстов производилась в разных районах Северной Осетии с 2007 по 2013 гг. Тексты записывались участниками экспедиций: О.И. Беляевым, А.П. Выдриным, Ю.В. Мазуровой, Н.В. Сердобольской и другими. Глоссирование текстов на иронском диалекте и их перевод на английский язык осуществлялось под руководством А.П. Выдрина участниками проекта: О.И. Беляевым, А.П. Выдриным и Ю.В. Мазуровой. Дигорские устные тексты глоссировались и переводились на английский язык по большей части А.П. Выдриным и Ю.В. Мазуровой. Глоссирование номеров дигорской газеты "Дигорæ" осуществлялось О.И. Беляевым, А.П. Выдриным и Ю.В. Мазуровой. Уал лагкомские тексты глоссировал и переводил на английский язык А.П. Выдрин. Оцифровка иронских и уаллагкомских аудиозаписей и их перевод на русский язык осуществлялась М.В. Дарчиевой. Расшифровка и перевод на русский язык дигорских записей производилась А.Ю. Халлаевой и З.А. Халлаевой
.
В настоящий момент поддержкой и улучшением Устного корпуса осетинского языка занимается А.П. Выдрин.

Транскрипция

Для записи текстов используется стандартная осетинская орфография, которая довольно адекватно отражает осетинскую фонологию и морфологию, однако соответствия между символами алфавита и их фонетическим значением не всегда тривиальны. С таблицей соответствий между этой орфографией, символами МФА и нашей рабочей транскрипцией, используемой в англоязычных статьях, можно ознакомиться здесь .

Программы

Для анализа текстов использовалась программа SIL FieldWorks Language Explorer (FLEx) . Все тексты, разобранные в рамках нашего проекта, доступны в трех форматах: проект FieldWorks (.fwbackup), PDF и OpenDocument (.odt). Проект FieldWorks открывается бесплатной программой SIL FieldWorks Language Explorer (FLEx) . Программа имеет разработанную систему поиска и возможность экспорта отдельных текстов в разные форматы. Для открытия файлов в формате OpenDocument можно использовать свободно распространяемый офисный пакет OpenOffice.org, доступный здесь . Последние версии Microsoft Office, начиная с Office 2007, также могут открывать файлы в формате .odt.

Гранты и благодарности

Создание Устного корпуса осетинского языка происходило при финансовой поддержке Программы Президиума РАН «Корпусная лингвистика» . Мы благодарны жителям Северной Осетии, у которых были записаны устные тексты, а также
коллективу радиостанции "Алания" и в особенности Л.В. Губаевой, автору программы

"Одноклассники", за предоставление нам звукозаписей радиопередач.

Любые замечания, в том числе сообщения об обнаруженных ошибках и неточностях, просьба отправлять на наш адрес: ossetic.studies@gmail.com .