

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ИНСТИТУТ ЛИНГВИСТИЧЕСКИХ ИССЛЕДОВАНИЙ

МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ
«ДОКУМЕНТИРОВАНИЕ
ЯЗЫКОВ И ДИАЛЕКТОВ КОРЕННЫХ
МАЛОЧИСЛЕННЫХ НАРОДОВ РОССИИ»

Тезисы докладов
международной научной конференции
Санкт-Петербург, 14–16 октября 2019 г.

Санкт-Петербург
2019

Международная конференция «Документирование языков и диалектов коренных малочисленных народов России».
Тезисы докладов международной научной конференции,
Санкт-Петербург 14–16 октября 2019 г. — СПб.: ИЛИ РАН,
2019. — 88 с.

Редакционная коллегия: В. В. Баранова, Е. В. Головки,
М. З. Муслимов, М. А. Овсянникова, С. А. Оскольская,
А. М. Певнов, Е. В. Перехвальская, М. Ю. Пупынина,
Ф. И. Рожанский, А. С. Сметина, А. А. Сюрюн,
А. Ю. Урманчиева.

Оригинал-макет подготовила А. С. Сметина.
Издание осуществлено при финансовой поддержке РФФИ,
грант № 17-29-09097
«Типология механизмов взаимодействия русского языка с языками
малочисленных народов России».

ISBN 978-5-6040926-0-6
doi: 10.30842/9785604092606

© Коллектив авторов, 2019
© ИЛИ РАН, 2019

СОДЕРЖАНИЕ

Т. Б. Агранат	
Новая парадигма и вечные ценности полевой лингвистики....	6
Т. А. Архангельский	
Диалектные явления в корпусах социальных сетей	8
А. В. Архипов, М. М. Брыкина, Б. Вагнер-Надь	
Проблемы расшифровки звуковых материалов из селькупского архива А. И. Кузьминой	10
А. Бугаева, Е. Ю. Груздева, М. Тэмина, Ю. Янхунен	
Документирование в постязыковой период: в поисках айнских корней на Нижнем Амуре.....	12
В. В. Воробьева, Я. Ю. Малькова, И. В. Новицкая	
Система числовых, посессивных и падежных маркеров в восточных диалектах хантыйского языка	15
Р. В. Гайдамашко, А. Н. Левичкин	
Опыт создания цифрового архива коми-пермяцких рукописей	17
Е. В. Головкин	
Что можно и чего нельзя извлечь из архивных материалов: алеутские записи В. И. Иохельсона на восковых валиках 1909–1910 гг.	20
Н. В. Дубровская	
В чём различие «сине-зеленых» систем цветоименований в уральских и тюркских языках?.....	21
А. В. Дыбо, В. С. Мальцева	
Нормализация диалектных текстов хакасского языка для автоматической обработки парсером	23
В. В. Дьячков, И. А. Хомченкова	
Использование текстовых корпусов для изучения русско-горномарийского переключения кодов.....	26
М. А. Егорова	
О семантике показателей ‘степени достоверности’, или эвиденциалов, в эвенкийском языке, по данным говоров Иркутской области	28

М. М. Зимин	
Фонетико-фонологическое описание ульчского языка по экспедиционным данным и существующие в нём лакуны	31
О. Н. Иванищева	
Неизвестные данные об орографической лексике кильдинского саамского языка (на материале архивов Мурманской области)	35
О. А. Казакевич	
От социолингвистического обследования к фиксации текстов и элицитации: специфика полевой работы в ситуации языкового сдвига	37
Е. К. Кривошапкина	
Звучащая речь ламунхинских эвенов	40
П. И. Ли, Т. В. Тимкин	
Проблема диалектной принадлежности полевых материалов по хантыйскому языку с реки Тромъеган	44
Н. В. Лукина, С. А. Попова	
О лингвистических материалах в архиве В. Н. Чернецова и в аннотациях Е. Шмидт	47
М. Д. Люблинская, Н. Н. Миронова	
Йоканьгские саами	50
М. З. Муслимов	
О методике сбора языкового материала по прибалтийско-финским языкам Ингерманландии	51
А. Нахимовский	
Язык русского крестьянства XX-го века в поисках архива ...	54
И. В. Недялков	
Документирование языка уйльта (45 лет поездке автора на Сахалин)	56
Г. А. Некрасова	
Полевые дневники экспедиций к коми-язьвинцам В. И. Лыткина	57
А. М. Певнов	
Особенности каритива в некоторых восточных диалектах эвенкийского языка	60

М. Ю. Пупынина	
Экспериментально-ориентированная документация	
как метод изучения многоязычия:	
исследование Нижней Колымы.....	62
Ф. И. Рожанский, Е. Б. Маркус	
Парадигматические классы глагола в сойкинском	
ижорском	65
Н. М. Стойнова	
Система императива в ульчском языке: полевые данные	67
А. Ю. Урманчиева	
Показатели каритива в лесном и тундровом энецком:	
сколько их?	71
N. Aralova, B. Pakendorf	
Documentation of Negidal: Navigating through the data.....	73
A. Filchenko	
(Re-)Defining the legacy: integration of the Tomsk language	
archive data into modern LDD projects on endangered	
Siberian indigenous languages	75
S. Fujishiro	
Profiling Dolgan before its standardization.....	77
G. Moroz, Ch. Naccarato, S. Verhees	
Variation in two dictionaries of Botlikh.....	79
E. Perekhval'skaya	
Spatial cases in Udihe: a quantitative analysis	80
O. Potanina, A. Filchenko	
Possession in Ob-Yenisei Languages:	
documentation-to-typology-to-documentation	83
T. Salminen	
On the relationship of documentation, dialectology,	
and standardization: materials from Tundra Nenets	85
T. Tuisk	
University of Tartu archives of Estonian dialects and kindred	
languages: materials of Uralic languages spoken in Russia	87

Новая парадигма и вечные ценности полевой лингвистики

Т. Б. Агранат
(ИЯз РАН, Москва)

А. Е. Кибрик ставит вопрос о соотношении документальных и теоретических знаний: что важнее, ценнее, фундаментальнее для науки? Ответ на этот вопрос в какой-то степени влияет на прагматические рекомендации: а) исследователю: к получению какого знания надо прежде всего стремиться, и б) обществу: получению каких знаний надо оказывать первоочередную поддержку. Эмпирические знания — не более чем база данных для собственно научной деятельности. Но жизненный цикл даже лучших образцов теоретической лингвистики неизмеримо короче жизненного цикла образцов лингвистики документальной. В работе делается вывод о том, что сильные и слабые стороны документальной и теоретической лингвистики дополнительно распределены, и их существование невозможно друг без друга. Более того, противопоставление документальной и теоретической лингвистики весьма относительно. Чисто документальных и чисто теоретических научных работ не существует. В действительности речь может идти о пропорциях между данными и теорией, о главной цели того или иного исследования [Кибрик 1997].

В другой своей работе А. Е. Кибрик пишет, что конечной целью описания неизвестного языка-объекта, естественно, является создание его научного описания [Кибрик 1992: 54].

В последние годы принято считать, что наиболее важной задачей при полевой работе является сбор корпусов текстов. Разумеется, получение реальных текстов, отражающих живой язык, невозможно переоценить, особенно, если они записаны от последних носителей. Записи образцов речи придавали большое значение и исследователи прошедших веков, у современных лингвистов есть большое преимущество — мы

располагаем новыми техническими возможностями, позволяющими сохранить для потомков аудио- и видеофиксацию естественной речи.

Еще одна задача — элицитация, вспомогательное средство для установления грамматических явлений, которые не удаётся отследить по текстам.

Вряд ли можно поспорить с тем, что сбор данных по языкам, находящимся под угрозой исчезновения, — задача безотлагательная. Учитывая скорость, с которой в настоящий момент исчезают языки, лингвисты должны в аварийном режиме записывать тексты, которые в ближайшее время записать будет уже невозможно. Получается, что в настоящий момент в приоритете сбор данных, видимо, для грядущей научной деятельности. Можно ли конечную цель отложить до лучших времен, точнее, до худших, когда собирать в поле будет уже нечего и можно будет сосредоточиться на научных описаниях языков-объектов? Учитывая продолжительность человеческой жизни, получается в таком случае, что научные описания будут делать следующие поколения лингвистов, используя чужой «полуфабрикат». Вопрос об адекватности таких описаний остаётся открытым. Часто бывает, что составители справочных изданий, описывая по различным источникам языки, с которыми сами непосредственно не работали, так искажают действительность, что исследователи «не узнают» свои языки. Поэтому представляется, что одни и те же люди должны собирать полевой материал и делать научные описания, по-прежнему относя последние к конечной цели документирования.

Литература

- Кибрик А. Е. 1992. Очерки по общим и прикладным вопросам языкознания. М.
- Кибрик А. Е. 1997. О соотношении документальных и теоретических знаний // Вестник МГУ. Серия «Филология». № 5. С. 103–107.

Диалектные явления в корпусах социальных сетей

Т. А. Архангельский

(Университет Гамбурга, Германия)

Проведение диалектологических исследований обычно требует существенных ресурсов (финансовых и временных), поскольку в большинстве случаев означает работу во множестве населённых пунктов. В докладе будет рассмотрен альтернативный и намного менее затратный способ изучения диалектных явлений — с помощью текстов, написанных в социальных сетях. Материалом для исследования послужили корпуса социальных сетей (ВКонтакте) удмуртского, коми-зырянского, эрзянского и лугового марийского языков объёмом 0,8–4 млн словоупотреблений. Благодаря наличию географических метаданных для некоторых авторов, а также меньшей степени соответствия литературному стандарту в текстах соцсетей по сравнению с другими письменными источниками, априори можно предположить, что из них можно было бы извлечь нетривиальные диалектологические данные. Однако есть несколько очевидных препятствий, затрудняющих такие исследования:

(а) сильное влияние литературной нормы и стремление многих авторов ей следовать;

(б) отсутствие географических метаданных у многих авторов и невозможность проверить правдивость метаданных для тех, у кого они указаны;

(в) малый объём данных.

Я покажу, что несмотря на эти проблемы, использовать корпуса соцсетей в диалектологии имеет смысл, хотя бы для получения предварительных данных. Я рассмотрю несколько примеров лексических и грамматических диалектных явлений в этих языках и покажу, насколько результаты, полученные с помощью исследования соцсетей, соответствуют результатам, полученным ранее традиционными методами (например,

опубликованным в диалектном атласе удмуртского языка [Насибуллин и др. 2009]). В случае, когда речь идёт об относительно частотных лексемах или явлениях, распределение вариантов восстанавливается из корпуса довольно хорошо, за исключением литературного варианта (который обычно встречается на всей территории). С падением частотности полезность корпусов предсказуемо падает. Так, распределение лексем со значением «подорожник», которых в удмуртских диалектах насчитывается два десятка и которые редко встречаются в текстах, восстановить оказалось невозможно даже в неполном и приблизительном виде. Тем не менее, несмотря на неполноту получаемых данных, их точность всегда высока: мне не удалось обнаружить существенных противоречий между данными, полученными из соцсетей и из традиционных источников. Дополнительная выборочная проверка точности указываемых авторами географических данных также показывает, что на них можно полагаться при идентификации диалектных явлений; при этом значение поля «место рождения» даёт намного более качественные результаты, чем «место проживания».

Литература

- Насибуллин Р. Ш., Максимов С. А., Семёнов В. Г., Отставнова Г. В. 2009. Диалектологический атлас удмуртского языка. Карты и комментарии. Вып. 1. Ижевск: НИЦ «Регулярная и хаотическая динамика».

Проблемы расшифровки звуковых материалов из селькупского архива А. И. Кузьминой¹

А. В. Архипов¹, М. М. Брыкина¹, Б. Вагнер-Надь^{II}

(^I – *Университет Гамбурга / МГУ;*

^{II} – *Университет Гамбурга, Германия)*

В Гамбургском университете с 2016 г. в рамках проекта INEL [Arkhipov, Däbritz 2018] продолжается изучение селькупских материалов А. И. Кузьминой (1924–2002), переданных ею в Институт финно-угроведения / уралистики в 2001 г.

Архив А. И. Кузьминой включает 30 рукописных томов полевых записей, собранных ею в период 1962–1977 гг. в различных селькупских диалектах, а также более 30 (оцифрованных) плёнок с аудиозаписями. Более подробное описание архива см. в [Тучкова, Хелимский 2010]. В данном докладе речь пойдёт только о записях текстов, поскольку именно они являются предметом изучения в проекте INEL. Первая версия корпуса селькупских текстов из рукописного архива опубликована в декабре 2018 г. [Brykina et al. 2018]. Выход второй версии запланирован на первую половину 2020 г.

Количество и качество материалов, представляющих различные селькупские диалекты, неоднородно [см. Тучкова, Хелимский 2010]. В рукописной части гораздо более полно представлены центральные и южные диалекты. Большая часть звуковых записей, наоборот, относится к северным диалектам. Качество звука варьирует в широких пределах, речь зачастую

¹ Доклад подготовлен в рамках исследовательской Программы Академий, совместно финансируемой Федеральным правительством Германии и Федеральными землями, при участии Федерального министерства образования и научных исследований и Вольного и ганзейского города Гамбурга. Программа Академий координируется Союзом Академий наук Германии.

неразборчива, и транскрибирование представляет собой чрезвычайно трудоёмкую задачу.

Благодаря компьютерным технологиям, ограниченность ресурсов для непосредственной полевой работы могла быть отчасти компенсирована за счёт самостоятельной работы информантов с последующим переслушиванием участниками проекта. В то же время транскрипции, выполненные носителем без участия исследователя, имеют некоторые существенные особенности.

Звуковые записи северноселькупских текстов составляют около 6 ч, из них расшифровано почти 5 ч: часть в ходе совместной работы с информантами (в 2017 и 2018 г.), прочие — самостоятельно Е. С. Сморгуновой. Из особенностей её записи можно отметить, например, неотражение вариативности окончаний *-n/-t* (GEN) и *-t/-tit* (3PL).

Бóльшие проблемы представляет расшифровка текстов на центральных и южных диалектах, где почти не осталось говорящих. Наиболее активным и компетентным информантом в настоящее время является И. А. Коробейникова. Было решено воспользоваться её помощью для расшифровки не только её родного (центрального) диалекта, но и южных. За исключением короткой совместной работы в Гамбурге (2016), расшифровка велась ею самостоятельно. В первую очередь были затранскрибированы записи из Напаса (центральный; около 25 мин). Значительно более представительный массив записей Ф. Ф. Тобольжиной (Иванкино / Тогур; среднеобской) — почти 2,5 ч — расшифрован И. А. Коробейниковой с помощью Н. П. Иженбиной. Наконец, из 3 ч записей из Усть-Озёрного (среднекетский) ими же расшифровано около 35 мин, причём для ряда текстов информанты работали независимо, что позволяет отметить различия между их записями (более точная транскрипция vs. более связный перевод); в целом транскрипции текстов на кетском диалекте селькупского языка информанты признают очень приблизительными.

Наконец, интерес представляют звукозаписи, расшифровки которых есть в рукописном архиве, которые также позволяют проследить вклад информанта в процесс транскрибирования.

Литература

- Тучкова Н. А., Хелимский Е. А. 2010. О материалах А. И. Кузьминой по селькупскому языку. Hamburg.
- Arkhipov C. L., Däbritz. 2018. Hamburg corpora for indigenous Northern Eurasian languages // *Tomsk Journal of Linguistics and Anthropology*. 3 (21). P. 9–18.
- Brykina M., Orlova S., Wagner-Nagy B. 2018. INEL Selkup Corpus. Version 0.1. Publication date 2018-12-31. Archived in Hamburger Zentrum für Sprachkorpora.
URL: <http://hdl.handle.net/11022/0000-0007-CAE5-3>.

Документирование в постязыковой период: в поисках айнских корней на Нижнем Амуре¹

А. Бугаева^I, Е. Ю. Груздева^{II}, М. Тэмина^{III}, Ю. Янхунен^{II}
(I – Токийский университет науки, Япония;
II – Хельсинкский университет, Финляндия;
III – Хабаровский краевой институт развития образования)

Район Нижнего Амура с древних времен является центром сосредоточения языков и культур, разнообразие которых сохраняется до сих пор. В этом регионе по-прежнему говорят на тунгусских языках (нанайском, ульчском, удэгейском, эвенкийском и негидальском), а также на амурском диалекте нивхского языка. Все перечисленные языки имеют

¹ Исследование выполнено в рамках проекта Japanese Government MEXT (Kakenhi) Grant-in-Aid for Scientific Research (C) #17K02743 (2017-2022) “Towards understanding dynamics of language change in Ainu” (PI: Anna Bugaeva).

грамматические описания, и, несмотря на угрозу их исчезновения, пока еще сохраняется возможность их документирования.

До недавнего времени на Нижнем Амуре говорили и на других тунгусских языках, а именно на орочском языке и языках кили (кур-урмийский) и килэн. Два последних языка обычно считаются вариантами нанайского языка, ср. [Суник 1958]. В рамках более подробной таксономии кили представляет собой смешанный эвенкийско-нанайский язык [Doerfer 1973], в то время как килэн является смешанным удэгейско-нанайским языком, последние носители которого (хэчжэ или хэчжэнь) живут в Китае [Kazama 1998]. Полевые исследования этих языков на территории России уже невозможны, однако все они так или иначе задокументированы.

Из различных этнографических и археологических источников известно, что в районе Амура проживали также айны, которые по разным причинам мигрировали сюда с о. Сахалин. Так, в конце XIX в. рядом с оз. Кизи существовало ульчское поселение Котон, название которого в переводе с айнского (*kotan*) означает 'селение'. Предполагается, что оно было некогда основано группой айнов [Смоляк 1975: 31]. Айны вступали в браки с представителями других амурских народов, входили в состав их родов, а иногда сами становились основателями новых смешанных родов. Например, в с. Кальма жил нивхский род, включавший в свой состав айнов и негидальцев [Штернберг 1933: 287]. Некоторые из современных ульчских родов, например Куйсали и Дуван, ведут свою историю от предков-айнов. В последние годы, во многом благодаря работам М. В. Осиповой [Осипова 2010], растет интерес к истории айнов на Нижнем Амуре.

Айнский язык, как известно, является одним из древнейших языков региона. На нём говорили на Хоккайдо, на Сахалине, на Курильских островах и на юге Камчатки [Дикова 1973]. В настоящее время все варианты айнского языка можно считать полностью утраченным. По айнскому

языку собрано большое количество материалов, однако, насколько нам известно, не существует ни лингвистической, ни социолингвистической информации о тех вариантах, на которых говорили на Нижнем Амуре и на Камчатке.

Летом 2018 г. авторы доклада провели экспедицию по Нижнему Амуре, одной из задач которой было собрать информацию об айнах и айнском языке. В современной языковой ситуации единственным методом, позволяющим получить какие-либо результаты, является сбор лингвистических биографий коренных жителей этого региона. В своем докладе мы хотели бы представить ту информацию, которую нам удалось собрать в ходе экспедиции, а также обсудить отдельные языковые проблемы, связанные с названием айнских родов на Нижнем Амуре в контексте общей проблематики постязыкового документирования.

Литература

- Дикова Т. М. 1983. Археология южной Камчатки в связи с проблемой расселения айнов. М.: Наука.
- Осипова М. В. 2010. Потомки айнов на Нижнем Амуре // Известия Института наследия Бронислава Пилсудского 14. С. 204–214.
- Смоляк А. В. 1975. Этнические процессы у народов Нижнего Амура и Сахалина. Середина XIX — начало XX в. М.: Наука.
- Суник О. П. 1958. Кур-урмийский диалект. Исследования и материалы по нанайскому языку. Л.: Государственное учебно-педагогическое издательство.
- Штернберг Л. Я. 1933. Гиляки, орочи, гольды, негидальцы, айны. Статьи и материалы / Я. П. Алькор (Кошкин) (ред. и предисл.). Хабаровск: Дальгиз.
- Doerfer G. 1973. Das Kur-urmiische und seine Verwandten // Zentralasiatische Studien. Bd. 7. S. 657–699.
- Kazama Sh. 1998. Genetical Position of Hezhen // Languages of the North Pacific Rim. Vol. 4 / O. Miyaoka, M. Oshima (eds). Kyoto: ELPR. P. 115–131.

Система числовых, посессивных и падежных маркеров в восточных диалектах хантыйского языка¹

В. В. Воробьева¹, Я. Ю. Малькова¹, И. В. Новицкая^{II}
(I – ТГУ / ТПУ; II – ТГУ, Томск)

В данном исследовании представлены результаты, полученные в ходе сравнительного анализа системы именных морфологических маркеров в четырёх диалектах языка восточных ханты (ваховском, васюганском, сургутском и салымском). Материалом исследования послужили грамматики восточных диалектов хантыйского языка, начиная с первых грамматических заметок, сделанных финскими учеными-филологами во время их путешествий по Сибири (М. А. Кастреном, К. Ф. Карьялайненом и Х. Паасоненом) в конце XIX — начале XX вв., до полных описательных грамматик отдельных восточных диалектов, основанных на материалах экспедиций текущего столетия [Karjalainen, Vértés 1964; Терешкин 1961; Gulya 1966; Honti 1984; Filchenko 2010; Csepregi 1998, 2017]. В исследование вошли также собственные записи, полученные в полевых условиях в 2017–2018 гг. от носителей ваховского диалекта, проживающих в пос. Ларьяк и с. Корлики Нижневартовского района Ханты-Мансийского АО.

Внимание было сосредоточено на маркерах, составляющих парадигмы трёх именных категорий: число, посессивность и падеж; исследование было нацелено на выявление общих и уникальных диалектных особенностей. Проведенный анализ

¹ Исследование выполнено при финансовой поддержке Министерства образования и науки Российской Федерации (грант № 14.Y26.31.0014) в рамках реализации научно-исследовательского проекта «Языковое и этнокультурное разнообразие Южной Сибири в синхронии и диахронии: взаимодействие языков и культур».

позволил систематизировать языковые данные по каждому восточному диалекту и выявить существующие между грамматическими описаниями противоречия, касающиеся терминологии, а также форм и наличия в тех или иных парадигмах именных морфологических показателей. Все выявленные формальные противоречия в падежной, числовой и посессивной системах приведены в таблицах. Наиболее дискуссионной является категория падежа, представленная в исследуемых диалектах разным количеством — от 7 до 11. В связи с этим, все выделяемые падежи были разделены на две группы: обязательную и факультативную. В первую группу попадают падежи, которые выделяются всеми хантологами: номинативный, агентивный-локативный, абессивный, лативный-иллативный-дативный-аллативный, инструментальный-комитативный, инструментальный-объектный и транслативный. Вторая группа включает падежи, выделяемые несколькими или только одним из исследователей: аблативный, аллативный, компаративный дистрибутивный и иксплетивный. Большинство падежей в языковой системе восточных ханты полифункциональны, что обусловило различия в терминологии. В качестве решения проблемы обозначения падежей было предложено использовать термины, отражающие основные функции падежных маркеров в восточных вариантах хантыйского языка. Например, маркер *-a/-ä* в ваховском диалекте передаёт значения латива, иллатива, датива и аллатива, по этой причине он обозначается сложным термином «лативный-иллативный-дативный-аллативный маркер».

Проведённое исследование именной системы восточных вариантов хантыйского языка позволило выявить уникальные диалектные особенности морфологических форм. В перспективе исследование морфологической системы имени существительного в восточных диалектах хантыйского языка будет дополнено анализом функционального аспекта числовых, посессивных и падежных маркеров.

Литература

- Терешкин Н. И. 1961. Очерки диалектов хантыйского языка (ваховский диалект). Л.: Наука.
- Чепреги М. 2017. Сургутский диалект хантыйского языка / Н. Б. Кошкарева (ред.). Ханты-Мансийск: Печатный мир.
- Csepregi M. 1998. Szurguti osztják chrestomathia. Studia uralo-altaica supplementa 6. Szeged: József Attila Tudományegyetem.
- Filchenko A. Yu. 2010. Aspects of the Grammar of Eastern Khanty. Tomsk: TSPU-Press.
- Gulya J. 1966. Eastern Ostyak Chrestomathy // Uralic and Altaic series 51. Indiana University. Bloomington — The Hague: Mouton.
- Honti L. 1984. Chrestomathia Ostiastica. Budapest: Tankönyvkiadó.
- Karjalainen K. F., Vértés E. (ed.). 1964. Grammatikalische Aufzeichnungen aus ostjakischen Mundarten // Memoires de la Societe Finno-Ougrienne 128. Helsinki.

Опыт создания цифрового архива коми-пермяцких рукописей¹

Р. В. Гайдамашко^I, А. Н. Левичкин^{II}

(I – ИЛИ РАН / РГПУ им. А. И. Герцена; II – ИЛИ РАН,
Санкт-Петербург)

Цель настоящего доклада — презентация проекта по созданию цифрового архива коми-пермяцких рукописей XVIII – начала XX вв.

Актуальность исследования определяется тем, что и поныне нет единого места хранения для всех пермяцких рукописей или хотя бы краткого справочника по их известным спискам. Даже количество дореволюционных рукописей, содержащих коми-пермяцкий языковой материал, в работах разных исследователей

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-012-00774а.

различается, и, возможно, цифра 60¹ не является окончательной. Кроме того, пермяцкие рукописи хранятся в различных архивах и библиотеках России, Венгрии, Германии, в большинстве своем оставаясь не введенными в научный оборот. Так, при работе по проекту издания первых коми-пермяцких словарей и первой коми-пермяцкой грамматики наш исследовательский коллектив столкнулся со сложностями в реконструкции генезиса, истории и типов взаимоотношений различных списков одного и того же документа. Цифровой архив, включающий фотокопии рукописей и их транскрипцию, был бы чрезвычайно полезен при текстологических и палеографических исследованиях на финно-угорском и более широком языковом поле.

Для достижения поставленной цели (онлайн представление оцифрованных рукописей) был создан сайт, находящийся сейчас на стадии разработки.

Для ввода, редактирования и представления текстов используется бесплатная, свободно распространяемая система управления контентом (CMS) Drupal. Такой подход позволяет снять разработку технических задач по программированию, разрешая использовать имеющиеся решения, которые предоставляет Drupal, а также эксплуатировать дополнительные модули (плагины) для расширения функциональности CMS. CMS Drupal также позволяет достаточно просто организовывать совместную работу по вводу и редактированию текстов.

В процессе работы были поставлены следующие технические задачи.

Корректность отображения пермской графики в машинописных текстах — решается использованием шрифтов стандарта Unicode, представленных на сайте и включающих необходимые глифы.

¹ 1917' Коми гижӧд 1920' воясӧдз.

URL: http://wiki.komikyv.org/index.php/Коми_гижӧд_1920'_воясӧдз
(дата последнего изменения: 17.07.2018).

Удобство пользовательского интерфейса с точки зрения анализа материалов. В большинстве случаев тексты различных списков представлены в табличном виде. Также реализовано подключение отображения графической информации (фотографий рукописей). Функционал основан на применении дополнительных сценариев на языке JavaScript.

Выбор средств отображения во многом зависит от способов сегментирования информации на сайте. Для лучшего восприятия рекомендуется разделение информации на небольшие части. В основе представления лексикографических материалов лежит словарная статья. Определённые структурные части словарных статей выделены как поля (заголовочное слово, варианты, толкование, перевод слова, варианты перевода, грамматическая информация и др.). Также заданы поля для классификации лексики по тематическому (идеографическому) принципу, для дополнительных грамматических характеристик и др. Присутствует возможность скрытия / отображения необходимых статей.

Помимо сегментирования информации, немаловажную роль играет её структурирование с точки зрения навигации. Перспективной в отношении цифрового представления рукописных языковых данных является связь словарных статей с корпусом текстов на основе лемматизации слов.

Задачей сайта является также отображение полного корпуса коми-пермяцких слов, отмеченных в известных рукописях, то есть каждое слово в материалах будет иметь лексическую и грамматическую характеристику.

Что можно и чего нельзя извлечь из архивных материалов: алеутские записи В. И. Иохельсона на восковых валиках 1909–1910 гг.¹

Е. В. Головки

(ИЛИ РАН / Европейский университет в Санкт-Петербурге)

Во время работы на Алеутских островах в 1909–1910 гг. русский этнограф Владимир Ильич Иохельсон записал на восковых валиках данные от носителей всех трёх алеутских диалектов. Для записи использовалась передовая техника того времени (фонограф). Иохельсон также сделал черновую транскрипцию большей части записанных текстов. Работу Иохельсона можно считать одной из самых ранних попыток профессионального документирования языковых данных. Почти все тексты были проверены и заново затранскрибированы в 1980-х гг. Кнудом Бергсландом и Мозесом Дирксом. По инициативе Майкла Краусса, директора Центра изучения языков коренного населения Университета штата Аляска в Фербенксе, тексты были опубликованы в 1990 г. [Jochelson 1990]. Однако несколько текстов на ныне не существующем диалекте острова Атту (последний носитель которого Джон Голодофф умер в 2008 г.) оставались незатранскрибированными и неопубликованными. В докладе будут представлены основные результаты работы по транскрибированию этих текстов, которая в 2008–2009 гг. (в рамках проекта U.S. National Science Foundation, руководитель Майкл Краусс) проводилась Мозесом Дирксом и автором настоящего доклада. В докладе будет представлена краткая история алеутских записей на восковых

¹ Исследование выполнено при поддержке гранта РФФИ 17-29-09097 «Типология механизмов взаимодействия русского языка с языками малочисленных народов России».

валиках, а также будет рассмотрена проблема транскрибирования исчезнувшего языка. Будут затронуты и более общие вопросы, в частности о существующих ограничениях на извлечение информации из старых записей. Эти ограничения определяются представлениями исследователей того времени о способах фиксации языкового материала и, в частности, сводятся к отсутствию контекста записей, специфике первоначального транскрибирования, отсутствием жанрового разнообразия.

Литература

Jochelson W. 1990. Unangam Uniikangin kayux Tunusangin — Unangam Uniikangis Ama Tunuzangis — Aleut Tales and Narratives / K. Bergsland and M. L. Dirks (eds). Fairbanks: University of Alaska, Alaska Native Language Center.

В чём различие «сине-зеленых» систем цветонаименований в уральских и тюркских языках?¹

Н. В. Дубровская
(ТГУ / ТГПУ, Томск)

Настоящее исследование вносит вклад в разработку проблем, связанных с вопросами лексической семантики и типологии систем цветонаименований. Рассматриваются причины эволюции систем с макроцветообозначениями; подробно анализируются особенности сочетаемости терминов цвета, кодирующих синюю, зеленую и желтую части спектра в диахронии и синхронии. Языковая репрезентация этой части

¹ Исследование выполнено при финансовой поддержке гранта Правительства РФ № 14.Y26.31.0014, проект «Языковое и этнокультурное разнообразие Южной Сибири в синхронии и диахронии: взаимодействие языков и культур».

спектра характеризуется использованием единого термина для сине-зеленой, желто-зеленой, или сине-желто-зеленой частей спектра в уральских и тюркских языках.

Благодаря возможности картографирования исследуемых явлений, проведённое исследование показало, что в XX в. для значительной части языков Евразии было характерно наличие сине-зеленых систем цветообозначений. Можно предположить, что причиной разрушения таких систем в языках европейской части материка послужило тесное взаимодействие с носителями иранских, волжских и пермских языков в Сельджукском и Булгарском государствах.

В Сибири, наоборот, тюркские и тунгусские языки с «сине-зеленой» системой оказали ассимилирующее влияние на носителей обско-угорских и самодийских языков, у которых в XVIII в. была зафиксирована «желто-зелено-синяя» система, сменившаяся к XX в. у большинства носителей «сине-зеленой». Эти изменения, реально зафиксированные в словарях и памятниках письменности, показывают, как происходит эволюция систем с макроцветообозначением. Анализ материала более чем ста языков и диалектов языков Евразии позволяет установить, что системы с макроцветообозначением образуют ареалы [Норманская 2019]. При интенсивных контактах носителей конкурирующих систем возможны два сценария развития:

1) Обе системы распадаются, макроцветообозначение исчезает, такая эволюция произошла в языках Малой Азии, в настоящее время в огузских, иранских, греческом языках макроцветообозначений нет.

2) Побеждает более устойчивая система: в Сибири это была «сине-зеленая» тюркская и тунгусская системы; менее устойчивая (обско-угорская, самодийская) проходит через этап тройного «макроцвета» и постепенно совпадает с более устойчивой.

В настоящее время на территории России наиболее распространённой является система русского языка, в которой нет макроцветообозначений. Как показывают наши полевые

исследования, в настоящее время в большинстве уральских, тюркских и тунгусских языков системы с макроцветообозначениями исчезли [Lingvodoc 2019].

Сбор и анализ типологических анкет, направленных на исследование сочетаемостных особенностей терминов цвета, позволяет отметить, что современные системы изучаемых миноритарных языков Сибири стремятся к появлению отдельных терминов для цветов спектра по модели европейских систем, в которых есть термины «синий», «желтый» и «зеленый». При этом, наблюдения над сочетаемостью цветоименований даёт возможность не только обнаружить типологически отличительные черты, но и объяснять их описанными выше историческими процессами.

Литература

- Норманская Ю. В. 2019. Карта по системам ударения, гласным фонемам и их изменениям, многозначности в уральских языках. Электронный ресурс (дата обращения: 25.05.2019).
Lingvodoc. 2019. Система электронных словарей и корпусов текстов уральских и алтайских языков.
URL: lingvodoc.ispras.ru (дата обращения: 25.05.2019).

Нормализация диалектных текстов хакасского языка для автоматической обработки парсером

А. В. Дыбо, В. С. Мальцева
(ИЯз РАН, Москва)

В докладе идёт речь о принципах представления диалектных текстов для их дальнейшей обработки автоматическим парсером, ориентированным на литературный хакасский язык. Парсер работает с электронным грамматическим словарем хакасского литературного языка, сделанным на основе [Субракова 2006] (22 тыс. статей). В словаре использована

стандартная орфографическая запись (кириллица с дополнительными знаками).

В данной ситуации важно соблюсти баланс между адекватным отражением диалектных особенностей и возможностью правильного анализа автоматическим парсером максимального количества словоформ. Для разрешения ситуации можно: а) редактировать словарь, внося в него новые единицы; б) вносить в грамматическую таблицу парсера новые аффиксы; в) изменять правила анализа уже введенных аффиксов, г) приближать представление диалектных словоформ к литературной норме. После парсера анализ словоформ правится человеком, поэтому возможно и д) давать парсеру литературный аналог, исправляя впоследствии представление словоформы на диалектное. При выборе варианта действий учитывается частотность конкретного явления.

Слой представления диалектного текста, поступающий на вход автоматической разметки, — это не фонетическая транскрипция, а «нормализация» диалектного текста, в которой используются только знаки хакасского алфавита.

Исходя из потребностей парсинга принимались следующие конкретные решения:

1. Анализ словоформ диалектного текста, которые имеют морфонологическое представление, совпадающее с литературным. (О морфонологическом представлении см. подробнее [Дыбо и др. 2019]).

1.0. Диалектный консонантизм при прочих равных чаще остаётся неприведённым к литературному аналогу, чем вокализм. Причина, прежде всего, в том, что орфографическая запись хакасских согласных ориентирована скорее на фонологическое представление, а гласных — на морфонологическое.

1.1. Из представления текста снимаются особенности фонетической реализации общего для разных диалектов морфонологического представления словоформ в тех случаях, когда количество реализаций морфоном совпадает. Примеры:

а) бельтырское и шорское [ĕ] соответствует лит. *i* [ɨ]: орфографически и в нормализации *пилім* (бельт. [pilĕm], лит. [pilɨm]) ‘моя спина’; б) *чол* (качинское [ʃ’ol], саг. [ʃ’ul], лит. [čol]) ‘дорога’; в) *парчадыр* (кач. [parčeadır], лит. [parčadīr]) ‘идёт’.

Точно таким же образом снимаются стяжения и выпадения, оглушения и озвончения, не соответствующие литературной норме, например: *парыбысты* (саг. [pa:ri[̇]sti]) ‘он ушёл’.

1.2. Если в диалекте реализации морфемы больше, чем в литературном языке, дополнительные реализации отражаются в представлении текста. Пример: в ряде говоров морфонема {Н} получает фонемную реализацию /d/ после морфонема {л} и {р}. Это бывает как на границе аффикса, так и в чередующихся основах, ср.: *ирди* 1) лит. [ирн-і] (‘губа-ЗРОС’) ‘его губа’, 2) лит. [ир-ні] (‘мужчина-АСС’) ‘мужчину’.

2. Анализ словоформ, имеющих отличный от литературного языка морфемный, в том числе морфонемный, состав.

2.1. Основы с отличными от литературных морфонемами вносятся в словарь. Пример: кызыльское, шорское *аш* ‘ячмень’ (лит. *ас*).

2.2. Дополнительные словоизменительные морфемы вносятся в грамматическую таблицу. Пример: саг. *парахча* ‘хочет идти’ содержит диалектный аффикс перспектива {АК}. Сюда же попадают диалектные аффиксы, совпадающие с литературными, но не подчиняющихся рядному сингармонизму, например, саг. {че} / кач. {ча} (лит. {чА}) ‘PRES’, саг., шор. {ох} ‘ASS’ (лит. {ОК}).

3. Менее регулярным образом отличающиеся лексические морфемы (основы и словообразовательные аффиксы) также заносятся в словарь, если встречаются в текстах часто. Примеры: саг., шор., бельт. *ник/них* (лит. *інек*) ‘корова’, *малты* (лит. *палты*) ‘топор’, бельт. *уннуң* ‘очень’ (лит. аналог неясен).

Схема расположения аффиксов, их семантика и правила реализации корректируются, исходя из проанализированных данных. В докладе будут более подробно рассмотрены разные случаи диалектных явлений, представляющие интерес для данной темы.

Литература

- Субракова О. В. (ред.). 2006. Хакасско-русский словарь. Новосибирск.
Дыбо А. В., Крылов Ф. С., Мальцева В. С., Шеймович А. В. 2019. Сегментные правила в автоматическом парсере Корпуса хакасского языка // Урало-алтайские исследования. 32(2).
Корпус хакасских диалектов:
URL: https://linghub.ru/oral_khakas_corpus/search

Использование текстовых корпусов для изучения русско-горномарийского переключения кодов¹

В. В. Дьячков^I, И. А. Хомченкова^{II}
(I – ИЯз РАН; II – МГУ / ИРЯ РАН, Москва)

Целью настоящей работы является обсуждение возможностей, которые предоставляют корпуса устных текстов при изучении переключения кодов. Это явление, наблюдаемое в том числе в финно-угорских языковых сообществах, привлекает в последнее время пристальное внимание исследователей (см., например, недавний сборник [Tánczos et al. 2018]). Вместе с тем, переключение кодов, используемое в горномарийской языковой среде, до недавних пор практически не становилось предметом отдельного исследования (см., впрочем, [Гаврилова 2013, 2014] о близкородственном луговом марийском языке).

Материалом для нашего исследования послужил, во-первых, находящийся в свободном доступе корпус горномарийского языка, созданный участниками проекта под руководством

¹ Исследование поддержано грантом РФФИ №18-312-00155 мол_а.

Е. В. Кашкина и насчитывающий 44297 токенов¹. Этот корпус был собран в ходе полевой работы в 2016–2018 гг. и включает в себя как диалогическую, так и монологическую речь. Во-вторых, для изучения структурных (синтаксических) характеристик переключения кодов в программной среде ELAN была создана небольшая коллекция текстов, состоящая из диалогов с большим количеством включений на русском языке. Мы подробно охарактеризуем разметку, созданную для этого корпуса, и опишем теоретические представления о переключении кодов, которые легли в её основу. В-третьих, в качестве материалов также были привлечены данные эlicitации и экспериментов. Большинство данных было получено в ходе нашей полевой работы в 2016–2019 гг. в Горномарийском районе Республики Марий Эл (сельское поселение Кузнецово).

В ходе нашего доклада на примере нескольких тематических исследований мы продемонстрируем, как материалы разного типа позволяют решать различные задачи, стоящие перед исследователями переключения кодов в горномарийском языке. Так, на материале основного корпуса мы покажем взаимосвязь между признаками, которые влияют на переключение кодов внутри количественной конструкции: синтаксический тип числительного, арифметическое значение и семантический тип контекста. Эти признаки во многом пересекаются, и, чтобы выявить их значимость, мы воспользуемся статистическими методами. Вместе с тем, большой корпус не всегда позволяет выявить значимые синтаксические характеристики переключения кодов, и поэтому текстовая коллекция, созданная в среде ELAN, была направлена в первую очередь на выявление именно этих характеристик. Исследование позволило получить ряд нетривиальных результатов: например, мы выявили, что в исследуемом говоре,

¹ Корпус доступен по адресу: <http://hillmari-exp.tilda.ws/corpus>.

по-видимому, практически не встречается переключения кодов внутри предложения, а большинство включаемых в речь составляющих на русском языке являются дискурсивными маркерами, адвербиалами, группами послелого или речевыми клише, что позволяет говорить о начальном этапе языкового сдвига. Наконец, полученная в ходе анализа текстовой коллекции информация послужила основой для экспериментов, в ходе которых мы протестировали некоторые ограничения на русскоязычные включения, о которых мы также подробно расскажем в нашем докладе.

Литература

- Гаврилова В. Г. 2013. Марийско-русское переключение и смешение кодов // Вестник Удмуртского университета. 2. Серия «История и филология».
- Гаврилова В. Г. 2014. Особенности внутрифразового марийско-русского кодового переключения // Финно-угорский мир. 2. С. 31–34.
- Tánczos O., Kovács M., Puura U. (eds.). 2018. Multilingual Practices in Finno-Ugric communities. Uralica Helsingiensia 13.

О семантике показателей ‘степени достоверности’, или эвиденциалов, в эвенкийском языке, по данным говоров Иркутской области

М. А. Егорова
(ИЯз РАН / РГГУ, Москва)

Согласно грамматическим описаниям, в эвенкийском зафиксирована система из трёх суффиксов, *-nĀ-*, *-rkA* и *-rgu*, которые описываются как показатели ‘наклонения вероятности’ в [Konstantinova 1964], ‘степени достоверности’ в [Nedjalkov 1997] и как показатели эвиденциальности с эпистемическим оттенком в [Bulatova, Grenoble 1999]. Три суффикса передают различные временные значения (настоящее / недавнее прошлое, прошлое и хабитуалис, соответственно).

Нашей целью было выяснить, используются ли названные показатели в речи современных носителей эвенкийского и — в случае положительного результата — какие значения они могут передавать.

В ходе экспедиции к эвенкам Иркутской области летом 2018 г. (токминско-верхнеленский и ербогачёвский диалекты), были собраны данные, полученные путём элицитации, и записи рассказов.

Согласно нашим данным, из трёх показателей информанты чаще всего используют *-nĀ-*; суффикс *-rkA* встретился нам только один раз (I).

- (I) *Lalbuka-l upkat d'əpə-v-čə-l.*
 ягель-PL весь есть-PASS-P.ANT-PL
 'Весь ягель съеден'.
Ə-dū ilə-l urīnčə-rkə-l.
 DEM-DAT человек-PL жить.на.стойбище-EP.PAST-3PL
 'Должно быть, тут была стоянка'.

Суффикс *-nĀ-* указывает на то, что сообщаемый факт расценивается как предположение, основанное на общих знаниях о мире (II), косвенных свидетельствах (III) или ощущениях говорящего (IV). Этот суффикс может усиливаться лексическими показателями эпистемической оценки (V).

- (II) *Motor d'av-dū āčīn. D'oromo-nō-ro.*
 мотор лодка-DAT нет украсть-EP-3PL
 'Мотора нет на месте. Должно быть, украли'.
 (III) *Gara kapurga-ra-n. Bəjūn ətə-nā-n.*
 Ветка хрустнуть-AOR-3SG лось идти-EP-3SG
 'Ветка хрустнула. Должно быть, лось прошёл'.
 (IV) *Bi agī-lī ŋənə-d'ə-Ø-m.*
 Я тайга-PROL идти-IMPF-PRES-1SG
 'Иду я по тайге'.

Plungian V. A. 2001. The place of Evidentiality within the universal grammatical space // *Journal of Pragmatics*.

Plungian V. A. 2011. Types of verbal Evidentiality marking: an overview // *Linguistic Realization of Evidentiality in European Languages*. Berlin — N.Y.

Фонетико-фонологическое описание ульчского языка по экспедиционным данным и существующие в нём лакуны

М. М. Зимин
(ИЯз РАН, Москва)

Ульчский язык принадлежит тунгусо-маньчжурской семье, континентальной группе нанайской (т. е. южно-тунгусской) ветви тунгусской подсемьи. Последние несколько десятилетий этот язык детально документируется специалистами разных дескриптивных школ (Саппоро — Токио — Киото, Санкт-Петербург, Москва), по записанным к настоящему моменту материалам можно составить достоверные суждения о фонетике и фонологии ульчского языка. В ряде случаев экспедиционные материалы значительно уточняют фонетико-фонологические модели классических публикаций [Петрова 1936; Суник 1985; Schmidt 1925]. Традиционной нанайстике было свойственно опираться на перцептивный анализ звучащего на языке текста, полученные данные публиковались без доказательства фонологичности отдельных сущностей (по структуралистской или какой-либо иной доказательной процедуре), фонетические инварианты одной лексемы зачастую включались в словари как отдельные входы без указания на сферу их употребления. Без обращения к полевым данным подобные описания языков не могут быть адекватно использованы для получения однозначных результатов; полученный при экспериментальном анализе набор фонологичных сущностей ульчского языка представляется следующим:

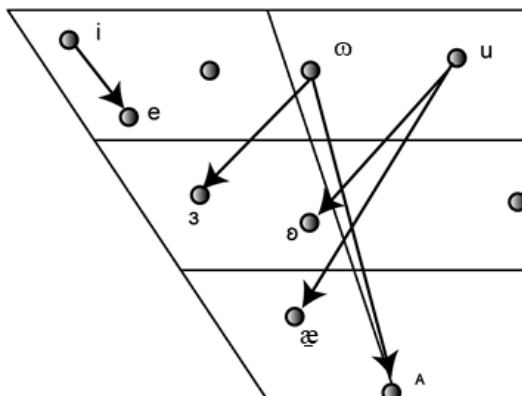
Таблица 1. Согласные

	bidental	labial	coronal				palatal	dorsal	radical	nasophar.	laryngeal						
			anterior	posterior	alveopal.												
central	nasals	ŋ	n	(ŋ)	ɲ	—	ŋ	ŋ̃	—	—	—						
	stops	p' ɸ	t' d'			ʃ	k' ɡ	p' ɓ	ɬ	ɮ							
	taps		r ~ r ~ l														
	trills																
	affricates	(q' ɸ)	—	tɕ'		(ɟ)	(k' ɡ)	(q' ɟ)	—	—							
fricatives	—	f (ɳ)	s ~ θ ~ ð	—	—	j	—	—	—	—							
glides	—	v	(s ~ θ ~ ð)	(r)	—	—	(v)	—	—	—							
lateral	affricates	—	l ~ ʎ ~ l					—	(l ~ ʎ ~ l)	—	—						
	fricatives	—															
	glides	—															
	taps, trills	—															

Таблица 2. Монофтонги

	front		front- central		central		central- back		back	
high	i iː	ɪ ʊː				ʊ ʊː		u uː		
high-mid	e eː	(ə)		ɜ ɜː		—		o oː		
						—				
mid-high						—				
						—				
mid	—					—				
	—					—				
mid-low	—	—	æ		ə		ɔ ɔː			
	—	—								
low-mid	—	—								
	—	ʌ ʌː								
low	—									

Рисунок 1. Дифтонги



Поскольку основная часть собранного в экспедициях материала представлена текстами, а корпус элициаций словоформ охватывает малое количество идиолектов, то часть фонетической системы не может быть на настоящий момент успешно фонологизована и соотнесена с классическими источниками. Лакуны в материале существуют по следующим аспектам фонетико-фонологической структуры:

- фонологический статус полудолгих гласных (*i̯*, *u̯*, *o̯*, *ʌ*);
- репрезентативный набор контекстов для ниже-средних долгих гласных (*ɛː*, *æː*);
- полный перечень всех долгих конечнослоговых согласных и их фонологический статус;
- детали распределения *ʌ* в собственной и заимствованной лексике;
- репрезентативный список контекстов, в которых встречается сегмент *ɟ*;
- перечень морфов, в которых встречается *ŋ* на месте графемы //н// словарей Т. И. Петровой и О. П. Суника;
- данные по хиатусу гласных (сочетания, в которых всегда наблюдается зияние vs. сочетания, в которых зияние лишь вариант поверхностной реализации фонологических комплексов V + glide + V);
- диалект-специфичные списки контекстов, в которых возможен сегмент *ð*;
- структурированные данные по диалект-специфичной фонотактике (инициальные и медиальные консонантные кластеры).

Указанные проблемы невозможно решить на настоящем этапе документации языка, необходимой является работа по сбору лексических данных (на основе упоминавшихся ранее словарей) и по сбору фонетических данных по сравнительно-историческому опроснику для нанийских языков. По предложенному алгоритму также возможно устранить существующий

недостаток данных по морфофонологии ульчских грамматических показателей, что является более экономичной альтернативой созданию ульчского корпуса глоссированных текстов.

Литература

- Петрова Т. И. 1936. Ульчский диалект нанайского языка. Л.
Суник О. П. 1985. Ульчский язык. Л.
Schmidt P. 1925. The language of the Olchas // Acta Universitatis Latviensis. Latvijas Universitātes raksti. 8. Rīga. P. 229–288.

Неизвестные данные об орографической лексике кильдинского саамского языка (на материале архивов Мурманской области)

О. Н. Иванищева
(МАГУ, Мурманск)

В докладе представлен анализ лингвистических данных, собранных краеведом и знатоком саамского языка Василием Кондратьевичем Алымовым (1883–1938). Цель доклада — показать важность использования архивных документов в работе по созданию тезауруса исчезающего языка. Краеведческий материал может послужить базой для сравнительных исследований, а также для выявления лакун в тематической группе орографической лексики (связанной с горами и возвышенностями, рельефом местности) кильдинского саамского языка. Материалы статей В. К. Алымова могут и должны быть критически осмыслены, но важность языковых данных, представленных в них, несомненна. Лингвистическая интуиция краеведа, его знание кольско-саамских языков, а также замечания по поводу специфики полевой работы делают эти немногочисленные работы важным источником.

В докладе анализируются неопубликованные архивные материалы 1930-х гг. из фонда Мурманского областного краеведческого музея.

В записке В. К. Алымова «О географических названиях на Кольском полуострове» [Алымов 1935а] от 20 июля 1935 г., сохранившейся в архиве Мурманского областного краеведческого музея, указывается на специфику пространственной ориентации у кольских саамов и приведен список географических терминов разных языков кольских саамов. В «Кратком словаре саамских физико-географических терминов, встречающихся на Кольском полуострове» В. К. Алымова [Алымов 1935b] от 9 июля 1935 г. представлены основные сведения о диалектах кольских саамов и словарь географических терминов, собранных Г. Д. Рихтером, взятых из словаря А. Генеца и собранных от саамского населения самим В. К. Алымовым, знавшим кильдинский саамский язык.

В записке «О географических названиях на Кольском полуострове» В. К. Алымов [Алымов 1935а] утверждает, что названия гор, озер и рек на Кольском полуострове, за небольшим исключением, идут от корней саамского языка. Причиной этому В. К. Алымов видит тесную связь саамов с природой. «Саамский язык — язык охотников, рыболовов и оленеводов. И как таковой, он богат и разнообразен физико-географическими терминами, названиями представителей фауны и флоры, с которыми саами издавна соприкасаются, названиями, относящимися к охоте, рыболовству и оленеводству» [Алымов 1935а: 1–10б].

«Краткий словарь саамских физико-географических терминов, встречающихся на Кольском полуострове» В. К. Алымов [Алымов 1935b] предваряет очерком о диалектном членении кольско-саамских языков. «Сводка», по выражению В. К. Алымова, географических терминов саамского языка представлена 61 позицией, охватывающей названия горного и водного ландшафта. Словарь интересен не только перечислением

географических терминов, но и выделением лексем разных кольско-саамских языков и диалектов. В «Кратком словаре саамских физико-географических терминов, встречающихся на Кольском полуострове» [Алымов 1935b] выделяются следующие тематические группы: названия разновидностей горных и водных объектов и их частей, водного пространства, места для поселения и ловли рыбы, каменистых и песчаных поверхностей, поверхностей с разными видами растительности, названия меньших по размеру объектов.

Сокращения

КП — книга поступлений;

МОМ — Мурманский областной музей.

Литература

Алымов В. К. 1935а. О географических названиях на Кольском полуострове. МОМ КП 15815/9. Л. 1–11.

Алымов В. К. 1935b. Краткий словарь саамских физико-географических терминов, встречающихся на Кольском полуострове. МОМ КП

От социолингвистического обследования к фиксации текстов и элицитации: специфика полевой работы в ситуации языкового сдвига¹

О. А. Казакевич

(ИЯз РАН / НИВЦ МГУ / ИЛ РГГУ, Москва)

Исходя из своего опыта полевой работы последних двух десятилетий в Сибири и на Дальнем Востоке, я предполагаю обсудить некоторые специфические черты работы лингвиста в дву- и многоязычных локальных сообществах, отказывающихся

¹ Доклад подготовлен в рамках проекта РНФ № 17-18-01649 «Динамика языковых контактов в циркумполярном регионе».

(или уже отказавшихся) от своего этнического языка в пользу языка более престижного и функционально более мощного. Сегодня в Сибири и на Дальнем Востоке этим более престижным языком, как правило, является русский, но исторический контекст показывает, что так было не всегда. В докладе будет рассмотрено соотношение трёх основополагающих методов работы полевого лингвиста — социолингвистического обследования локального сообщества, фиксации текстов и элицитации — в зависимости от степени изученности современного состояния локальных вариантов языков, представленных в сообществе, где работает лингвист, и стадии языкового сдвига.

Очевидно, что для успешной работы в ситуации языкового сдвига лингвисту необходимо адекватно представлять языковую ситуацию в сообществе, количество носителей каждого из языков сообщества, их распределение по возрастам и уровень владения каждым из языков в разных возрастных группах, поэтому без социолингвистического обследования целевого сообщества сегодня не обойтись. Что касается сбора лингвистических данных, то полезно придерживаться некоторых общих стандартов их фиксации:

- для всех лингвистических данных, собираемых в поле — как для текстов, так и для анкет, — необходима аудиозапись; для фиксации текстов желательна также видеосъёмка;
- в каждом локальном сообществе стоит работать с как можно большим количеством информантов разного возраста, чтобы в полученных данных были зафиксированы языковые варианты каждого из поколений, в которых язык еще сохраняется;
- интересные данные можно получить не только от информантов, хорошо владеющих языком, но и от тех, чье владение языком не идеально, и даже от тех, для кого язык остаётся лишь воспоминанием.

Тексты — это полевые данные ‘первого порядка’, самые желанные для лингвиста, особенно если речь идёт о неопи-санном или слабо описанном идиоме, находящемся на продвинутой стадии развития языкового сдвига в этно-локальной группе ср. [Dixon 2007]; практически все остальные типы данных в принципе могут быть извлечены из текстов при достаточном объёме корпуса и наличии аудиозаписи корпуса. Более того, тексты содержат также ценнейшие экстра-лингвистические данные об истории, традиционной культуре, инновациях и просто о повседневной жизни группы. Однако не следует забывать, что сам сбор текстов включает элицитацию как один из важных элементов формирования корпуса, а именно, — расшифровку текстов с помощью информантов-носителей языка.

Элицитация — неизбежная часть полевой работы, поскольку корпус текстов языка, особенно языка минори-тарного, никогда не бывает достаточно полным. Среди прочего¹, она чрезвычайно полезна для проверки результатов анализа корпуса. В докладе будут приведены примеры взаимного дополнения анализа текстов и элицитации, а также роли информации о языковой ситуации в сообществах для интерпретации полученных лингвистических данных.

Литература

- Chelliah S. L., De Reuse W. J. 2011. Handbook of Descriptive Linguistic Fieldwork. Dordrecht: Springer.
- Dixon R. M. W. 2007. Field Linguistics: A Minor Manual. Sprachtypologie und Universalienforschung (Focus on Linguistic Fieldwork) / A. Yu. Aikhenvald (ed.). 60 (1). P. 12–31.

¹ Разные типы элицитации подробно описаны в [Chelliah, De Reuse 2011].

Звучащая речь ламунхинских эвенов

Е. К. Кривошапкина
(БГПУ, Благовещенск)

В современном мире происходит унификация национальных различий, поэтому язык и самосознание малочисленного народа остаётся важным инструментом трансляции культурного наследия следующим поколениям. Язык ламунхинских эвенов В. И. Цинциус относит к западному диалекту эвенского языка [Цинциус 1947: 270], а К. А. Новикова — к крайне-западному наречию [Новикова 1958: 205].

Село Себян-Кюель (Кобяйский район Республики Саха (Якутия)) основано в 1932 г. По данным переписи 2010 г., население составляет 800 человек, из них мужчин — 324, женщин — 217, детей — 255: учащихся школы — 170, в ДОУ — 45 воспитанников. Представителей малочисленных народностей — 685 [Кривошапкина 2012: 18]. Для выяснения уровня владения языком среди ламунхинских эвенов и перспектив сохранения и развития их языка и культуры мы побеседовали с жителями, посетили культурные мероприятия, провели анкетирование среди учащихся 5-х–11-х классов МБОУ «Себян-Кюельская национальная эвенская средняя общеобразовательная школа» им. П. А. Ламутского.

Опрос показал, что 30 учащихся (41%) из 73 считают, что хорошо знакомы с традициями; 40 учеников (54,7%) отмечают, что слабо знакомы с культурой своего народа. Из национальных традиций учащиеся перечислили гостеприимство и *нимат* (дарение добычи гостю — мяса и шкуры добытого животного); уважительное отношение к старшим и пожилым; кормление огня на новом месте; почитание духа огня, горы, реки; запрет кричать на новом стойбище; запрет играть с огнём, чтобы не растерять оленей; необходимость перешагивать через упряжь),

назвали традиционные праздники эвенского народа (*Оралчимна инэун, Эвинэк, Хонначан инэун* и др.).

Респонденты должны были перечислить пути сохранения эвенского языка. Среди предложенных были ответы:

- родители должны сами говорить с детьми и между собой на родном языке;
- надо изучать родной язык и развивать оленеводство;
- проводить мероприятия на эвенском языке;
- издавать на родном языке книги, журналы;
- создавать языковую среду в школе и проводить дополнительные занятия по родному языку.

51 (69,8%) респондент из 73 указывают на проблему утраты родного языка.

Уровень знания якутского и русского языков 42 ученика (57,5%) оценивают высоко, хотя есть сомнения в их владении якутским. Респонденты указали, что дома разговаривают на эвенском и якутском языках (по 46,5% — 34 ученика), и лишь 6,8% (5 учеников) разговаривают на русском языке. С друзьями и знакомыми ученики общаются «одинаково на русском, якутском, эвенском» — 34 (46,5%). Свой уровень владения родным языком считают: очень хорошим — 13 учащихся (17,8%), хорошим — 21 (28,7%), средним — 31 (42,4%), ниже среднего — 4 (5,4%) и низким — 4 ученика (5,4%). Наибольшими трудностями при обучении родному языку ученики считают: изучение правил орфографии — 37 человек (50,6%), 13 детей считает трудным чтение (17,8%) и 11 (15%) — письмо. 52 респондента (71,2%) отмечают, что родной язык в школе преподаётся на среднем уровне, 11 (15%) считают уровень обучения высоким, а 10 учеников (13,6%) — низким. По мнению учащихся, это зависит, прежде всего, от отсутствия или недостаточного количества учебников по грамматике и орфографии эвенского языка отдельно с 5 по 11 классы (34 опрошенных — 46,5%).

Другой источник — анкетирование взрослых. Анкету заполнили 83 человека в возрасте от 20 до 83 лет. По результатам анкетирования 68 человек (81,9%) признают, что в современном обществе возрастает актуальность проблемы незнания родного языка; 72 человека (86,7%) отмечают, что в настоящее время ощущается острый недостаток учебных пособий, книг, онлайн-текстов и СМИ на эвенском языке; в школе, кроме книг местных авторов (А. В. Кривошапкина, В. С. Кейметинова — Баргачана, Д. В. Кривошапкина — Нимкалана), отсутствуют учебники по грамматике эвенского языка и художественная литература.

Для сохранения и распространения эвенского языка в наслеге, районе, республике респонденты предлагают вести передачи местного телевидения и радио на родном языке и другие пути.

78 человек (93,8%) оценивают знание родного языка как хорошее и удовлетворительное. У большинства в домашней библиотеке имеются книги эвенских писателей и поэтов. 68 респондентов (81,9%) считают необходимым сохранение традиционного быта и уклада кочевников, 77 человек (92,7%) отмечают важность наличия кочевых семей. В настоящее время в государственном унитарном предприятии «Себян» и родовых общинах круглогодично кочуют 15 семей.

Деятельность промышленных компаний на территории Ламынхинского наслега (АО «Прогноз» на 2-х участках «Вертикальная» и «Хогин», ООО «Чочимбал», ООО «Кристалл», АО «РосГео») имеют лицензии на разведку и добычу серебра и золота одобряют 32 опрошенных (38,5%), не одобряют 36 респондентов (43,3%). Те, кто одобряет, считают, что они содействуют социально-экономическому развитию наслега; те, кто не одобряет, обеспокоены ухудшением экологии и окружающей среды, состоянием здоровья населения и нарушением традиционного природопользования и уклада жизни КМНС.

Таким образом, можно сделать вывод, что жители Ламынхинского наслега обеспокоены будущим родного языка и предлагают пути дальнейшего развития. Как отмечает В. А. Кейметинов, эвены сами должны «проявлять активную заботу о возрождении, сохранении и развитии родного языка» [Кейметинов 2013: 10]. В Себян-Кюеле дети считают эвенский язык родным, знают его, но большинство ребят дома, в школе и со знакомыми и друзьями в повседневной жизни общается на якутском языке. Некоторые родители между собой говорят на родном языке, а с ребёнком — на якутском. Интервьюируемые отмечают, что нет учебников по эвенской грамматике с 5-го по 11-й классы, что на уроке родного языка надо говорить только на эвенском языке.

Литература

- Кейметинов В. А. 2013. Эвенско-русский словарь. Толкование и этимология — Эвэди-ньучиди төрэрүк. Төрөн йак дьүгүлин үкчэнүттин ньан онут идук һиедэнгэн. Ч. 5 (доп.). Якутск: Компания «Дани-Алмас».
- Кривошапкина Е. К. 2012. Три сестры: сказки, стихи и статьи — Илан нёнул асаткар: нимкалал, дёнтурал. Якутск: Бичик.
- Новикова К. А. 1958. Основные особенности эвенских говоров Якутской АССР // Докл. и сообщ. ИЯз АН СССР. Вып. 11. М. С. 186–205
- Цинциус В. И. 1947. Очерк грамматики эвенского (ламутского) языка. Л.

Проблема диалектной принадлежности полевых материалов по хантыйскому языку с реки Тромъеган¹

П. И. Ли^I, Т. В. Тимкин^{II}

(I – НГУ / Институт филологии СО РАН;

II – Институт филологии СО РАН)

Одной из причин неоднозначного выделения говоров сургутского диалекта хантыйского языка является то, что сургутское *jǎŋ* ‘река’ передаётся на русском как: *аган*, *юган*, *еган*, *ёган*. Притоки Оби, по которым называют говоры сургутского диалекта, носят имена Большой Юган, Малый Юган, Аган, Тромъеган, название последнего особенно вариативно. В «Чертёжной книге Сибири» С. У. Ремезова [Ремезов 2007] название реки — *Торь Юганъ*. Позднее на картах употребляются другие обозначения реки: *Тром-Юган*, *Тромъеган*. Сейчас на картах используется только вариант *Тромъеган* (*Тромъёган*), хотя названия *Тром-Юган*, *Тром-Аган* есть в водном реестре и в официальных новостях.

Тромъеганский диалект называют различно: *Tremjuga*n [Karjalainen, Toivonen 1948], *тром-юганский* [Терёшкин 1981], *тромъеганский* [Чепреги 2017]. В некоторых работах наряду с тромъеганским говором выделяется тромаганский, но *Тром-Аган* — название поселка, а не реки. Выделяя *тремюганский*, *тромаганский* [Хонти 1993: 302], Л. Хонти говорит о говорах поселка и реки: *die Mundart im Tromagan, die Mundart am Tremjuga*n [Honti 1981: 106]. В работе [Abondolo 1998: 362] это различие считается не географическим, а историческим (современный *тромаганский* диалект на месте *тромъеганского*).

¹ Исследование выполнено при финансовой поддержке РФФИ (проект 19-012-00388 «Специфика звукового строя сургутского диалекта хантыйского языка»).

Мы обработали фонетические данные от семи информантов: с р. Лямин, из пос. Русскинская, пос. Тром-Аган, пос. Ортъягун (данные полевых экспедиций ИФЛ СО РАН 2014–2018 гг.). С помощью акустического анализа в программе Praat мы обнаруживаем фонетические различия в области краткого вокализма. Фонема /ɔ/ (орф. *ö*) у информанта с Лямина отсутствует, у информантов из Ортъягуна произносится закрыто, возможно, нейтрализуется с /o/. Фонема /ɯ/ (орф. *ü*) реализуется у информантов различно: как заднерядная гласная, как среднерядная гласная, как среднерядная гласная с делабиализацией передних оттенков. Обозначим зафиксированную вариативность на карте (Рис. 1).

Более важной изоглоссой представляется наличие фонемы /ɔ/. По этому признаку информанты с Тромъёгана и Ортъягуна могут быть объединены как носители одного — тромъеганского — говора, информант с Лямина относится к другому говору, вероятно, пимскому. Тромъеганский говор неоднороден. Возможно, следует говорить не о противопоставлении тромъеганского и тром-аганского говоров, а о необходимости выделять подговоры в составе одного говора.

Литература

- Ремезов С. У. 2007 [1701]. Чертежная книга Сибири: факс. воспроизведение рукописи Рос. нац. б-ки. Тобольск: Возрождение Тобольска.
URL: <https://archive.org/details/RemezovCHertezhnayaKnigaSibiri/page/n2> (дата обращения: 31.05.2019).
- Терёшкин Н. И. 1981. Словарь восточно-хантыйских диалектов. Л.: Наука.
- Хонти Л. 1993. Хантыйский язык // Языки мира. Уральские языки / В. Н. Ярцева (ред.). М.: Наука. С. 300–319.
- Чепреги М. 2017. Сургутский диалект хантыйского языка / Н. Б. Кошкарева (ред.), А. С. Песикова (ред. хантыйского текста),

Рисунок 1. Зафиксированные идиомы: 1 — р. Лямин (1 информант), 2 — пос. Русскинская (2 информанта), 3 — пос. Тром-Аган (1 информант), 4 — пос. Ортъгун (3 информанта).



Обозначения

Фонема /ɔ/:

- перешла в /ɛ/,
- ◆ выделяется,
- ◆ сближается с /o/.

Фонема /u/:

- ◆ среднерядное произношение с делясализацией передних оттенков,
- среднерядной произношение
- ◆ заднерядное произношение.

- Т. А. Ефремова (пер.). Ханты-Мансийск: Печатный мир г. Ханты-Мансийск.
- Abondolo D. 1998. Khanty // The Uralic Languages / Abondolo D. (ed.). London — N.Y.: Routledge. P. 358–387.
- Honti L. 1981. Zur Frage nach der Herausbildung der Ostostjakischen Mundarten im Lichte der Lautgeschichte // Acta Linguistica Hungarica. T. XXXI. Budapest: Akadémiai Kiadó.
- Karjalainen K. F., Toivonen Y. H. 1948. Ostjakisches Wörterbuch. Lexica Societatis Fenno-Ugricae X/I–II. Helsinki.

**О лингвистических материалах в архиве
В. Н. Чернецова и в аннотациях Е. Шмидт¹**

Н. В. Лукина^I, С. А. Попова^{II}

(^I – *Научно-исследовательская лаборатория
«Музей и культурное наследие» ТГУ, Томск;*

^{II} – *Обско-угорский институт прикладных исследований и
разработок, Ханты-Мансийск)*

Научный архив В. Н. Чернецова поступил в Музей археологии и этнографии Сибири Томского университета в 1980 г. Это фонд № 869, состоящий из 157 единиц хранения (дел), содержащий полевые записи Чернецова 1920-х–1940-х гг. Этнографические материалы по ненцам, хантам и манси содержат образцы лексики этих языков и другие лингвистические материалы. Лексика этнографических записей частично представлена в публикациях Н. В. Лукиной и О. М. Рындиной [1987, 1999].

Фольклорные тексты на мансийском языке составляют около 1,5 тысяч листов (16 тетрадей, номера 39–54 в архиве) и представляют варианты языка разных групп манси, в частности

¹ Публикация подготовлена при финансовой поддержке РНФ, проект № 19-18-00329 «Интерпретация языковой и культурной истории народа манси: этнографические, фольклорные и лингвистические материалы архива В. Н. Чернецова».

с р. Конда-Тап, Сосьва, Ляпин, Обь (Вежакары, Ильпи-пауль), Ивдельский р-н Свердловской обл. (Анжух-пауль). Несколько тетрадей с записями фольклора в 1990-х гг. скопировала Е. К. Скрибник, их анализ дан в статье [Скрибник 1999]. Некоторые тексты использованы при подготовке электронного словаря-тезауруса [Thesaurus / Onomasiological Dictionary].

В 1984–85 гг. венгерский фольклорист Ева Шмидт подготовила аннотации к 175 текстам из 13 архивных дел (№ 39–47 и 51–54). Они изложены на русском языке, с частичным сохранением мансийской лексики. Сейчас рукописная версия Е. Шмидт и её машинописный вариант хранятся в Обско-угорском институте прикладных исследований и разработок (Ханты-Мансийск). В аннотациях Е. Шмидт есть сведения о разных видах транскрипции, применяемой Чернецовым: «В делах № 39–43 тексты записаны на основе Единого северного алфавита, разработанного в 1929–1930 гг. Начиная с дела № 44, Чернецов пользуется смешанным вариантом фонетической транскрипции *Setälä* и Единого северного алфавита (ЕСА). Долгота гласных поставлена нерегулярно и не всегда точно + на литературной транскрипции латинскими буквами + упрощённая фонетическая транскрипция на латинице». Е. Шмидт указывает страницы или места в тексте, содержащие языковые объяснения; имеются также её пометки такого рода: «образец букв фонетической транскрипции», «три мансийских предложения с русским переводом в полевой транскрипции В. Н. Чернецова» и др.

Мансийский язык в аннотациях представлен в небольшом объёме — это паспортизация текста, личные имена фольклорных персонажей, некоторые непереводимые слова.

При паспортизации текста Е. Шмидт приводит фамилию, имя или прозвище, реже отчество информанта; название текста; жанр и место записи. В некоторых записях информанты не указаны. Большая часть записанных фамилий имеет неверное звучание, например, Амендъев (должно быть Аментьев),

Курыков (Куриков), Пукшинов (Пуксиков), Щеков (Секов) и др. Есть фамилии на латинской графике Sampiltal (Самбиндалов), здесь же указано его прозвище Тор ојка. Личные имена фольклорных героев представлены в разных грамматических формах, но в большей части записаны на латинской графике, что составляет трудность в их прочтении; следовательно, для анализа текстов нужна предварительная транслитерация. Кроме того, для приведения записей в соответствие с современной мансийской грамматикой требуется расстановка долгот, которая почти везде отсутствует.

Место записи указано на мансийском языке. Жанр текстов почти везде написан на русском языке. Названия текстов не всегда имеют адекватные переводы, например, один из вариантов мифа о сотворении Земли «*Магыг тормыг сосхатыглам эрыг*» букв. 'Песня [как] Земля с Неба перелилась' в аннотации переведена: «Песенный миф об отоплении неба и земли».

Для исследования фольклорного материала в архиве В. Н. Чернецова требуется знание разных диалектов мансийского языка и умение работать с разными видами записей текстов.

Литература

- Лукина Н. В., Рындина О. М. (ред.) 1987. Источники по этнографии Западной Сибири. Томск.
- Скрибник Е. К. 1999. О мансийских языковых материалах из архива В. Н. Чернецова // Языки коренных народов Сибири. Вып. 5. Новосибирск: Изд-во СО РАН. С. 293–302.
- Lukina N. V., Ryndina O. M. 1999. Die Materialien von V. N. Tzernecov zur Mansischen Verwandtschaftsterminologie. (Finnisch-ugrische Mitteilungen 21/22). Hamburg.
- Thesaurus / Onomasiological Dictionary (corpus-based) // Лингвистический раздел: е-Словарь.
URL: <http://babel.gwi.uni-muenchen.de>

Йоканьгские саами

М. Д. Люблинская^I, Н. Н. Миронова^{II}

(^I – *ИЛИ РАН, Санкт-Петербург;*

^{II} – *пос. Умба, Мурманская обл.*)

Саами проживают в России на Кольском полуострове. Сейчас из четырёх саамских диалектов или языков осталось три: кильдинский (п. Ловозеро), самый большой по числу носителей (несколько сотен человек, разные источники приводят разные цифры, из них регулярно используют его не больше ста человек), колта-саамский (нотозерский в России, примерно 20 говорящих) и йоканьгский (терский) саамский, у которого от 2 до 10 носителей. Последняя носительница четвертого диалекта Кольского полуострова — бабинского — скончалась в 2003 г. [Саамские языки].

Нина Николаевна Миронова, урождённая Данилова, родилась в деревне Йоканьга в 1936 г. В 1960 г. закончила РГПИ и всю жизнь работала в п. Умба учителем немецкого языка. Хотя по неудачному стечению жизненных обстоятельств ей не удалось стать аспиранткой Г. М. Керта, она начала самостоятельно записывать материал о разных сторонах жизни йоканьгских саами, в том числе о йоканьгском варианте саамского языка. Поскольку этот язык является бесписьменным, слова записывались кириллицей по аналогии с русскими. Составлен небольшой словарь йоканьгского саамского, к которому сейчас делаем аудиозаписи, и разговорник на йоканьгском варианте по аналогии с кильдинским с записанным вариантом произношения фраз (в программе PRAAT).

В архиве хранится рукопись еще одного очень интересного словаря, собранного неизвестным исследователем-краеведом по фамилии Кривоногов, напечатанное на дореволюционной машинке, на котором помечено «Терский саамский».

Помимо материалов о языке в архиве Нины Николаевны собраны интересные этнографические наблюдения и комментарии к опубликованным работам о саамах, а также представлено описание истории рода Даниловых.

Литература

Саамские языки. URL: https://ru.wikipedia.org/wiki/Саамские_языки (дата обращения: 01.06.2019).

О методике сбора языкового материала по прибалтийско-финским языкам Ингерманландии¹

М. З. Муслимов
(ИЛИ РАН, Санкт-Петербург)

Все прибалтийско-финские языки Ингерманландии в настоящее время находятся в состоянии далеко зашедшего языкового сдвига и для речи практически всех носителей этих языков в настоящее время характерна та или иная степень аттриции. Это создаёт определённые трудности как при диалектологических исследованиях, так и при изучении грамматических систем этих идиомов. В докладе мы рассмотрим основные проблемы, встающие перед исследователем идиомов Ингерманландии и предложим возможные способы их решения.

Аттриция может проявляться на разных уровнях. На лексическом уровне она может проявляться либо в забывании информантом каких-то лексем, либо в заимствовании их из литературного финского, эстонского или другого говора. На морфологическом уровне возможны выравнивание парадигм по аналогии, утрата отдельных грамматических категорий,

¹ Исследование выполнено при поддержке гранта РФФИ 17-29-09097 «Типология механизмов взаимодействия русского языка с языками малочисленных народов России».

нейтрализация грамматических оппозиций. На синтаксическом уровне возможно калькирование синтаксиса доминирующих языков. На фонетическом уровне аттриция может проявляться, например, в утрате сингармонизма или в утрате противопоставления по долготе. При этом такого рода аттриционные изменения в идиолекте данного информанта в ряде случаев могут порождать словоформы, совпадающие с корректными, неаттриционными словоформами другого ингерманландского идиома, например словоформа *muissan* ‘помнить:1SGPRS’ в финском говоре Каприо (при *muistaa* ‘помнить:3SGPRS’ с другой ступенью чередования) в случае выравнивания по аналогии или влияния литературного финского языка может быть заменена формой *muistan*, которая представлена в большинстве финских диалектов Ингерманландии. Аналогичные эффекты возможны и на уровне лексики (например, дистрибуция вариантов *uks* и *ov(i)* ‘дверь’) и синтаксиса (колебания между генитивом и адессивом в конструкции *X-GEN/ADESS pittää* ‘X-у надо’).

Для отграничения случаев аттриции, заимствований из другого идиома, смешения кодов от корректных форм мы придерживаемся следующих общих принципов:

1. Каждая изоглосса или грамматическое явление должно рассматриваться по отдельности. Это означает, что исходя из имеющегося в нашем распоряжении материала мы можем предположить, что для части изоглосс у нас просто не наберётся достаточно материала, чтобы с уверенностью отделить аттриционные варианты от корректных. Это же касается и грамматики — если для одних участков системы мы в состоянии описать её ареальные вариации, то для других участков мы можем только констатировать наличие вариаций в принципе, но не сможем дать им ареальную интерпретацию.

2. Необходима как можно более густая сеть пунктов, откуда может быть получен языковой материал, в идеале она должна охватывать все населённые пункты, где есть или были в прошлом носители данного языка. Аттриционные варианты

в благоприятных случаях не образуют компактного ареала, а рассеяны отдельными вкраплениями по всему ареалу данного языка.

3. Необходимо привлечение как можно большего числа информантов, в том числе и с невысокой степенью владения языком. В настоящее время даже в идиолектах информантов с самой высокой степенью владения языком могут спорадически встречаться аттриционные и заимствованные формы; с другой стороны, даже информанты со степенью владения языком 5 и 6 по шкале Вахтина могут помнить такую лексику, которая у более компетентных информантов была вытеснена заимствованиями из литературного языка. Именно такова ситуация в говорах прихода Тюрё с вариантами названий картофеля. *tuna* (зафиксированного у большинства информантов, в том числе и от тех, кто помнил только небольшое количество слов) и *peruna* (зафиксированного у нескольких информантов со сравнительно высокой степенью владения языком, однако знакомых также и с литературным финским языком).

4. И элицитация, и записи спонтанной речи могут содержать как аттриционные варианты, так и корректные. При элицитации информанты могут испытывать затруднения при порождении редко употребляемых словоформ, которые тем не менее могут быть спонтанно порождены в речи. С другой стороны, в речи информантов, которые в настоящее время часто используют литературный язык, могут встречаться литературные варианты даже в том случае, когда они стремятся говорить по-ингерманландски. В этом случае элицитация может помочь выявить диалектные варианты.

В докладе будут рассмотрены конкретные примеры, иллюстрирующие эти принципы.

Язык русского крестьянства XX-го века в поисках архива¹

А. Нахимовский

(Университет Колгейт, Гамильтон, США)

История русского крестьянства в XX-ом веке была одной из запретных тем в советское время. Неудивительно, что, когда положение изменилось, историки, социологи, лингвисты, этнографы начали посылать экспедиции в деревни записывать жизненные истории крестьян, которые родились до коллективизации, а может быть, даже и до «германской войны». (Лингвисты-диалектологи, особенно Л. Л. Касаткин, начали такие записи еще раньше, но хранили их в архивах [Касаткин 2010]). Даже старшеклассники элитных гимназий в Петербурге и Москве ездили в летние экспедиции собирать этнографический материал, который часто, по сути, был историко-социологическим. Большую роль в начале этой работы сыграли Теодор Шанин и Виктор Данилов. Начиная с 1992 г. они начали посылать группы аспирантов, а то и просто энтузиастов, в разные концы страны записывать крестьянские истории, которые были собраны в книге [Ковалев 1996].

В начале XX в. в России было 100 миллионов крестьян, и никто не думал о них, как об исчезающей малой народности. Однако, начиная с середины 30-х гг. и особенно после войны, традиционная крестьянская речь перестала передаваться от родителей детям, см. немногочисленные работы [Дьячок 2003; Daniel et al., to appear; Nakhimovsky 2019]. Касаткин [1993] отмечает роль школьного образования в этом процессе. Как только язык или диалект перестает передаваться детям,

¹ Автор благодарит Уэльза Брауна, Ольгу Йокояма, Леонида Касаткина и Теодора Шанина за комментарии и поддержку в написании этих тезисов. Я один отвечаю за высказанные в них мнения.

спустя несколько десятилетий он становится исчезающим языком, и так и произошло после 1990-х с русским крестьянством.

В полевых работах конца XX в. историки и социологи пытались сохранить историческую память, а диалектологи продолжали собирать диалекты, которые в совокупности составляли исчезающий язык предвоенных и предреволюционных крестьян. Язык крестьян в целом не рассматривался как исчезающий и не документировался. Практически это означает, что не установлено стандартов записи и расшифровки, принципов публикации, организации и хранения цифровых архивов. Архивы рассыпаны по частным коллекциям и вузовским фонотекам, в которых не всегда царит полный порядок. Задача настоящего доклада — обосновать необходимость организации архива записей 1990–2015 гг., который сохранил бы более уже невоспроизводимый язык русского крестьянства XX-го века.

Литература

- Дьячок М. Т. 2003 Русское просторечие как социолингвистическое явление // Русский филологический портал. Гуманитарные науки. Вып. 21. С. 102–113.
- Касаткин Л. Л. 1993. Русские диалекты и языковая политика // Русская речь. 2. 1993. С. 82–90.
- Касаткин Л. Л. (ред). 2010. Русская деревня в рассказах её жителей. М.: АСТ-ПРЕСС.
- Ковалев Е. М. (ред.) 1996. Голоса крестьян. М.: Аспект Пресс.
- Daniel M., Kazakova P, Ter-Avanesova A. (to appear). Dialect loss in the Russian North: modelling change across variables // Language Variation and Change.
- Nakhimovsky A. 2019. The Language of Russian Peasants in the XXth Century. Lexington Books.

Документирование языка уйльта (45 лет поездке автора на Сахалин)

И. В. Недялков

(СПбГУ / ИЛИ РАН, Санкт-Петербург)

Подготовка к полугодичной экспедиции на о. Сахалин в 1974 г. (с мая по октябрь) к орокам (как их часто называли и называют в специальной литературе) началась с осени-зимы 1973 г. Подготовительная работа включала ряд этапов:

1) работу в библиотеке ЛГУ с кандидатской диссертацией Т. И. Петровой и её монографией [Петрова 1967]; составление по материалам Т. И. Петровой картотеки словаря уйльта;

2) чтение материалов К. А. Новиковой по её собственному предложению и с её помощью (помню её общую тетрадь в коленкоровой обложке с записанными карандашом 25 орокскими текстами с переводами и замечательными карандашными рисунками оленей и чумов);

3) изучение в Отделе алтайских языков статей Дз. Икегами по языку уйльта, которые он присылал О. П. Сунику, К. А. Новиковой, а впоследствии и мне.

Немаловажную роль после окончания факультета английского и немецкого языков играла работа с О. П. Суником по нанайскому и маньчжурскому языкам (чтение и разбор текстов).

Работа в экспедиции на о. Сахалин проходила в двух местах: с. Вал и г. Поронайск. В с. Вал особое внимание мне уделили носители северного диалекта языка уйльта Ольга Николаевна Семенова, её дочь Наталья Васильевна Семенова (которая оказала мне неоценимую помощь при анализе материалов Т. И. Петровой), а также Семен Александрович Надеин, который знал несколько языков народов той местности (как минимум, эвенкийский и уйльта). В г. Поронайск

важнейшую помощь как в записях текстов, так и в их расшифровке оказали носители южного диалекта уйльта Накагава Пакта, Накагава Анна, Накагава Василий, Сирюко (Людмила Хомовна) Минато. Эти люди относились ко мне с максимальной добротой и участием. Именно им я во многом обязан полученными материалами.

В результате работы, проведённой в экспедиции, мной было записано 10 текстов на северном диалекте и 27 текстов на южном диалекте (часть из них расшифрована; полной расшифровке помешали внешние обстоятельства); проверен текст книги Т. И. Петровой, выявлен ряд неточностей и требующихся исправлений; выявлен ряд технических трудностей экспедиционной работы. В список задач дальнейшей работы входит расшифровка собственных и других записей по языку уйльта, включая глоссировку и грамматические комментарии (материалы К. А. Новиковой и Т. И. Петровой у автора имеются); уточнение номенклатуры глагольных форм уйльта и подготовка собранных автором образцов уйльтинской речи (несколько сотен конструкций) к публикации.

Литература

Петрова Т. И. 1967. Язык ороков (уйльта) / О. П. Суник, В. И. Цинциус (ред.). Л.: Наука.

Полевые дневники экспедиций к коми-язьвинцам В. И. Лыткина

Г. А. Некрасова

*(Институт языка, литературы и истории
Коми научного центра УрО РАН, Сыктывкар)*

Коми-язьвинцы проживают в Красновишерском районе Пермского края по среднему и верхнему течению р. Язьва. По Всероссийской переписи населения 2010 г., в «Алфавитном

перечне возможных вариантов ответов населения для кодирования ответа на вопрос 7 Переписного листа формы Л» коми-язьвинцами записалось 803 человека [Алфавитный перечень]. Большинство исследователей сходятся во мнении, что язык коми-язьвинцев наряду с коми-зырянским и коми-пермяцким является отдельным наречием коми языка [Лыткин 1961; Лыткин, Тепляшина 1976: 106; КЯЭ 1998: 194, 199]. Р. М. Баталова считает его диалектом коми-пермяцкого языка [Баталова 1982: 3]. Основанием выделения языка коми-язьвинцев в отдельное наречие явилось наличие особых гласных фонем, не свойственных ни коми-зырянским, ни коми-пермяцким диалектам.

Первое описание коми-язьвинского наречия принадлежит Арвиду Генецу, который в 1889 г. изучал говоры сс. Верх-Язьва и Паршаково [Genetz 1897]. Ему удалось зафиксировать основные фонетически и грамматические особенности этого наречия. В 1949–1953 гг. изучением языка коми-язьвинцев занимался В. И. Лыткин, который совместно с С. А. Поповым совершил три экспедиции к коми-язьвинцам. Во время полевых экспедиций осуществлялся сбор как лингвистического, так и социолингвистического материала.

В Научном архиве Коми НЦ УрО РАН хранятся шесть полевых дневников экспедиций к коми-язьвинцам. Нельзя однозначно утверждать, что это все дневники всех трёх экспедиций, но на основе анализа даже одного дневника можно определить методы, использованные при изучении языка; сложности, которые возникали при записи и обработке материала, ценность собранного материала.

В полевой лингвистике традиционно применяются два основных метода работы с информантами: метод документирования (запись текстов) и метод элицитации (опрос информантов, переводы). Содержащийся в дневниках материал, включающий отдельные слова, словосочетания, предложения и тексты, свидетельствуют о том, что В. И. Лыткин использовал оба метода. Сбор лексического материала осуществлялся

по лексико-семантическим группам. В некоторых случаях для уточнения семантики слова сделаны рисунки. Текстовые материалы составляют фольклорные произведения (сказки, частушки, загадки, детские песни), рассказы производственного, бытового, автобиографического содержания, переводы произведений русских классиков. Записи произведены в фонематической транскрипции на кириллической основе с использованием латинских букв *i* и *j*, спорадически *d*, особых графем для передачи звонких аффрикат (*ʒ* и *ʒ̣*) и гласных фонем: *ə* для гласного заднего ряда ниже-среднего подъема, *ы* со знаком огубленности (*ɵ*) для огубленного гласного средне-переднего ряда верхнего подъема и *õ* для огубленного гласного средне-переднего ряда среднего подъема. В каждом слове отмечено ударение. Основную сложность при исследовании языка составило определение качества некоторых гласных, о чём свидетельствуют исправления букв, постановка знака вопроса над буквой.

Результатом исследования языка коми-язьвинцев стала монография [Лыткин 1961], но большинство собранного В. И. Лыткиным текстового материала осталось неопубликованным.

Литература

- Алфавитный перечень. URL: <https://gigabaza.ru/doc/80192.html>.
- Баталова Р. М. 1982. Ареальные исследования по восточным финно-угорским языкам (коми языки). М.: Наука.
- КЯЭ (Коми язык: энциклопедия). 1998. / Г. В. Федунёва (отв. ред.). М.: ДИК.
- Лыткин В. И. 1949. Полевой дневник экспедиции к коми-язьвинцам. Тетрадь 1 // Научный архив Коми НЦ УрО РАН. Ф. 21, оп. 1, ед. хр. 15.
- Лыткин В. И. 1961. Коми-язьвинский диалект. М.: Изд-во АН СССР.
- Лыткин В. И., Тепляшина Т. И. 1976. Пермские языки // Основы финно-угорского языкознания. Марийский, пермские и угорские языки. М.: Наука. С. 97–228.

Особенности каритива в некоторых восточных диалектах эвенкийского языка

А. М. Певнов

(ИЛИ РАН, Санкт-Петербург)

В докладе рассматриваются две формы каритива ('без кого / чего') в восточных диалектах эвенкийского языка.

Первая представляет собой результат превращения аналитической формы каритива в синтетическую: *гирки-жэ āчин* (друг-N.CNG отсутствие) > *гирки-жэчин* 'без друга' (N.CNG — именной коннегатив). Такая трансформация произошла в некоторых эвенкийских диалектах, вероятно, под влиянием якутского языка.

Формы эвенкийского каритива индифферентны по отношению к категории «человек / не человек», при том что противоположное комитативное значение ('с кем / чем') выражается разными аффиксами: если речь идёт о человеке, то употребляется аффикс *-нун* (*-н'ун*), если же имеется в виду предмет или животное, то используется показатель *-лкй̄н* (*-чй̄*).

Большой интерес представляют формы каритива личных местоимений. Их краткое описание в говорах эвенков Амурской области дано в работе [Булатова 1987: 45–46], где приведены следующие примеры: *минэ* ~ *миниуйэ āчин* 'без меня', *синэ* ~ *синиуйэ āчин* 'без тебя', *нууана* ~ *нууаниуйа āчин* 'без него', *митйэ* ~ *митиуйэ āчир* 'без нас' (инкл.), *мунэ* ~ *муниуйэ āчир* 'без нас' (экскл.), *нууара* ~ *нууартиуйа āчир* 'без них'. Н. Я. Булатова отмечает, что вариант с *минэ*, *синэ* и т. д., как более поздний, употребляют «представители среднего и молодого поколения эвенков». Любопытно, что в этих конструкциях также может происходить «слияние местоимения с отрицанием *āчин*, причём начальный гласный в слове *āчин*

подчиняется закону гармонии гласных: *миниуйэчин* ‘без меня’, *синиуйэчин* ‘без тебя’ и т. п.» [Булатова 1987: 46].

В 1981 г. М. М. Хасановой и мной были записаны от Н. И. Фёдоровой (отец её из рода Бута) в дер. Владимировка (р-н им. Полины Осипенко Хабаровского края) следующие конструкции с каритивным значением: *нууаун'а-дā āчин* ‘без него’, *нууаун'аһал āчин* ‘без них’, *муниун'э āчин* ‘без нас’, *суниун'э āчин* ‘без вас’ (две последние были предложены носителю мной). Подобные формы естественны для существительных. Каритивные формы эвенкийских личных местоимений представлены только в работе Н. Я. Булатовой. Записанные нами словоформы включают показатель множественного числа *-һал* (интересно, что именно он присоединяется к аффиксу *-лкāн*, выражающему значение обладания предметом); показатель именного коннегativa *-н'а/-н'э* (< *-ја/-јэ*); показатель отчуждаемой принадлежности *-у-* (возможность его присоединения к личным местоимениям заслуживает особого внимания); сегмент *-и-* можно считать соединительным гласным, однако не исключено, что это реликт показателя пратунгусо-маньчжурского родительного падежа; *мун-*, *сун-* — основы косвенных падежей личных местоимений, *нууа-* — основа местоимений 3-го лица.

Семантически близкими являются такие аналитические конструкции в эвенкийском языке: *бй āчин-дў-в* (я отсутствие-DAT.LOC-POSS.1SG) ‘в моё отсутствие’, *нууан āчин-дў-н* (он(а) отсутствие-DAT.LOC-POSS.3SG) ‘в его / её отсутствие’ и т. д.

Каритивные формы личных местоимений наподобие *нууаун'а āчин* ‘без него’ мне никогда не встречались в эвенкийских текстах.

Что касается других тунгусо-маньчжурских языков, то у меня нет сведений о том, как в них выражается каритив личных местоимений и существует ли он вообще. Если нет, то интересно знать, каким образом восполняется такая лакуна.

Литература

Булатова Н. Я. 1987. Говоры эвенков Амурской области. Л.: Наука.

Экспериментально-ориентированная документация как метод изучения многоязычия: исследование Нижней Колымы¹

М. Ю. Пупынина
(ИЛИ РАН, Санкт-Петербург)

В 2019 г. была проведена экспедиция по экспериментальному исследованию языкового сдвига в языках Якутии и Чукотки под влиянием русского и/или якутского языков. Участники проекта работали в каждом из этих регионов. Помимо документации, включающей сбор базовой социолингвистической информации и запись текстов, мы давали информантам ряд однотипных заданий, что облегчало сравнительный анализ идиолектов. Экспериментально-ориентированная документация в нашем проекте включает в себя четыре типа заданий.

- (1) Элицитация с использованием картинок, представляющих различные ситуации; справа от картинок даны подписи, содержащие один или два глагола в форме инфинитива / императива и одно или два существительных в форме абсолютива (в зависимости от типа ситуации). Очевидные минусы этого задания — необходимость грамотности и некоторой привычки к чтению на родном языке, а также

¹ Исследование выполнено в рамках проекта РНФ #17-18-01649 «Динамика языковых контактов в циркумполярном регионе», данные собирались в экспедициях по указанному проекту и по проекту NSF #1761551 “Investigating language contact and shift through experimentally-oriented documentation”.

наличие литературной нормы, не учитывающей диалектное варьирование.

- (2) Пересказ анимационного фильма непосредственно после просмотра.
- (3) Рассказ на основе четырёх связанных по сюжету рисунков.
- (4) Перевод (или название по рисунку) лексических единиц по темам «части тела» и «животные».

Целью исследования было выявить контактные изменения в структуре языка, касающиеся, главным образом, порядка слов, падежной системы, а также выражения пространственных отношений. Для этого опрашивались носители разных возрастов. Ожидалось, что более молодые информанты в среднем будут порождать высказывания, более подверженные влиянию русской (или якутской) грамматики, чем пожилые.

Работа велась на западе Чукотки, в Билибинском р-не, и на востоке Якутии, где пересекаются традиционные территории эвенов, чукчей и юкагиров. В этих районах можно встретить людей, владеющих, помимо русского, чукотским и эвенским языком (на Чукотке), или тремя и даже четырьмя языками: чукотским, эвенским, якутским и юкагирским (в Якутии). В рамках другого проекта по изучению языковых контактов в Арктике, в Нижнеколымском районе Якутии проводилось подробное социолингвистическое анкетирование, посвящённое многоязычию. Среди прочего, информантам надо было указать уровень своего владения каждым из языков региона по специальной семиуровневой шкале (от «совершенно свободно говорю» до «не знаю языка»). Оказалось, что задания (1)–(4) помогают сопоставить уровни владения различными языками и соотнести эти данные с собственной оценкой респондента. Так, было опрошено двое многоязычных людей, владеющих в различной степени всеми языками региона. В 2018 г. у них же были собраны социолингвистические анкеты.

Таблица 1. Данные анкетирования (2018)

	Информант 1	Информант 2
Русский	Совершенно свободно говорю, основной язык	Совершенно свободно говорю, основной язык
Чукотский	Совершенно свободно говорю, основной язык	Совершенно свободно говорю, основной язык
Эвенский	Говорю, но другие замечали ошибки	Объясняюсь простыми фразами и все понимаю
Юкагирский	Говорю, но другие замечали ошибки	Совершенно свободно говорю, основной язык
Якутский	Говорю, но другие замечали ошибки	Объясняюсь простыми фразами и все понимаю

Таблица 2. Данные экспериментально-ориентированной документации (2019). Результаты выполнения заданий (выполнено/не выполнено)

	Информант 1 (задание 1 / 2 / 3 / 4)	Информант 2 (задание 1 / 2 / 3 / 4)
Чукотский	Да / Да / Да / Да	Да / Да / Да / Да
Эвенский	Нет / Нет / Нет / Да	Нет / Нет / Нет / Нет
Юкагирский	Да / Да / Да / Да	Да / Да / Да / Да
Якутский	Да / Да / Да / Да	Нет / Нет / Нет / Нет

Информант 1 выше всего оценил свою компетенцию в чукотском, а владение юкагирским, эвенским и якутским оценил как среднее («говорю с ошибками»). Чукотские задания были выполнены им без затруднений, однако ни одного задания, связанного с эвенским языком, кроме (4), он выполнить не смог. При этом задания на якутском и юкагирском языках не вызвали у него серьезных затруднений. Информант 2 одинаково оценил свою компетенцию в чукотском и юкагирском и задания выполнял достаточно уверенно для обоих языков, а задания на эвенском и якутском выполнить не смог. Задания, разные по сложности выполнения, позволили уточнить оценку респондентом собственных знаний. Нужно

учитывать, однако, что неспособность информанта выполнить задания не говорит о том, что он вообще не может говорить на языке, и тем более ни одно из предложенных заданий ничего не говорит о рецептивных навыках.

Таким образом, методика, сочетающая социолингвистический опрос с выполнением разноплановых заданий по каждому языку, эффективна для изучения многоязычия. Но данная методика должна быть дополнена заданиями, фиксирующими рецептивные способности информанта по каждому языку.

Парадигматические классы глагола в сойкинском ижорском

Ф. И. Рожанский^I, Е. Б. Маркус^{II}

(I – *ИЛИ РАН, Санкт-Петербург / Тартуский университет;*

II – *Тартуский университет, Эстония*)

Исследование посвящено глагольному формообразованию в сойкинском диалекте ижорского языка — одного из малых прибалтийско-финских языков. К настоящему времени получение новых данных по ижорскому языку затруднено: из немногих существующих носителей уже почти никто не может работать языковым консультантом (в силу состояния здоровья или иных причин). Однако документация ижорского языка, проводимая нами с 2006 г., позволила собрать достаточное количество материала для описания основных явлений фонетики и грамматики.

В настоящем исследовании рассматриваются парадигмы 450 ижорских глаголов. Для большинства глаголов парадигмы записывались от нескольких носителей языка при помощи метода эликации. Задачей исследования было построение системы парадигматических классов.

С точки зрения глагольных грамматических категорий, ижорский язык не выглядит экзотически на европейском фоне: в нём есть категории времени (настояще-будущее, прошедшее,

перфект и плюсквамперфект), наклонения (индикатив, императив, кондиционалис), полярности, числа и лица (кроме личных форм существуют имперсональные формы). Но в отличие от языков средневропейского стандарта, где парадигматические классы глаголов определяются в первую очередь наборами разных словоизменительных суффиксов, а чередования в основе оказываются прерогативой ограниченной группы «неправильных» глаголов, в ижорском языке чередования в основе наблюдаются у большинства глаголов. Специфичный для сойкинско-ижорского тройной контраст согласных по длительности еще больше усложняет систему глагольных основ.

В работе [Saar 2017] для ижорского языка было выделено 113 глагольных парадигматических классов. В нашем исследовании мы предлагаем альтернативную систему, организованную иерархически. На первом уровне глаголы классифицируются по матричному принципу на основе трёх признаков: (1) одноосновные vs. двухосновные, (2) с показателем прошедшего времени *-i* vs. *-iži*, (3) не содержащие консонантных чередований vs. с чередованием ступеней vs. с геминацией vs. с чередованием ступеней и геминацией. Из 16 теоретически возможных классов в нашем материале встретились 12. На втором уровне выделяются подклассы: глаголы с дополнительным слогом в основе, глаголы с двойным чередованием, глаголы с дополнительной геминацией, односложные глаголы. На третьем уровне различаются подклассы, специфичные для отдельных классов и связанные с мелкими морфонологическими отличиями (как правило, это продление, сокращение или вытеснение гласных на стыке основы и показателя).

Предлагаемая система парадигматических классов может быть использована как при написании грамматики, так и при построении систем морфологического синтеза.

Литература

Saar E. 2017. Isuri keele Soikkola murde sõnamuutmissüsteem. PhD thesis. University of Tartu.

Система императива в ульчском языке: полевые данные¹

Н. М. Стойнова
(ИРЯ РАН / ИЯз РАН, Москва)

В докладе будет описана система императивных форм в ульчском языке (тунгусо-маньчжурские) по данным элицитации и текстов, собранных в экспедициях в Ульчский район Хабаровского края (2017–2018 гг.). Эти данные несколько отличаются от представленных в существующих описаниях. В докладе будут продемонстрированы основные отличия и высказаны предположения об их возможных причинах.

Набор императивных форм в ульчском языке включает формы всех лиц и чисел: собственно императивные (два набора форм 2 л. ед. и мн. ч.), гортативные (инклюзивные формы 1 л. мн. ч., ‘давай(те) сделаем (говорящий совместно с адресатом)’), императив 1 л. ед. ч. (‘давай я сделаю’) и юссивные формы (формы 3 л. обоих чисел, ‘пусть он / они сделает / сделают’), см. Табл. 1 и примеры (1)–(4)².

В докладе будет обсуждаться устройство ульчской императивной парадигмы с точки зрения типологии императивных систем, а также формальные и семантические особенности отдельных форм, интересные в более широком типологическом контексте.

Краткое описание императивных форм (примерно одинаковое) даётся в очерках [Петрова 1936: 57; Суник 1985: 45],

¹ Исследование выполнено в рамках проекта РНФ No. 17-18-01649.

² В целях экономии места ниже приводится и обсуждается только парадигма утвердительных форм.

см. Табл. 1. Полевые данные не вполне ему соответствуют: отличие касается форм 1 л. Нет форм на *-ži*, симметричных формам 3 л. (как в описаниях и в близкородственном нанайском, [Аврорин 1961: 123]): форма 1 л. мн. ч. полностью отсутствует, форма 1 л. ед. ч. содержит вместо маркера *-ži* презентный (1). Таким образом, теряется противопоставление по клюзивности ('мы, включая / исключая адресата'), ожидаемое по данным описания.

Табл. 1. Парадигма императива, глагол 'идти':
полевые данные vs. [Петрова 1936: 57]; [Суник 1985]¹

	SG	PL
1 excl	ηən-i-tə [ηənə-ži-tə]	Ø [ηənə-ži-pu]
1 incl	—	ηən-i-su
2	ηənə-ru	ηənə-ru-ksu
2 (nonspec)	ηənə-səri	ηənə-sər-tu
3	ηənə-ži-ni	ηənə-ži-ti

Охарактеризуем ульчскую парадигму (в виде, представленном в наших данных) с точки зрения внутренней однородности по процедуре, предложенной в [Гусев 2013: 223–244]². Формы однородны (различаются только суффиксом лица-числа) внутри каждого из лиц, границы подпарадигм проходят между всеми лицами, в т. ч. симметрично устроены формы 1 л. ед. ч. и 1 л. мн. ч. инкл. (Табл. 2). По данным [Гусев 2013], это второй по частотности тип императивных парадигм в языках мира (27%)³.

¹ Формы из предшествующих описаний даны в квадратных скобках. Без скобок — наблюдаемые. Цветом выделены клетки парадигмы, в которых они засвидетельствованы.

² Ср. также менее детальную процедуру оценки однородности по отношению к форме 2SG в работах [Аувера и др. 2004; Auwera et al. 2013].

³ Подсчёты для 2 и 3 л. и для 1 л. мн. ч. инкл. Поведение формы 1 л. ед. ч. рассмотрено в цитируемой работе как отдельный параметр варьирования.

Табл. 2. Однородные формы внутри парадигмы императива

	Полевые данные		Петрова / Суник	
	SG	PL	SG	PL
1 excl		Ø		
1 incl	—		—	
2				
3				

Еще один параметр — степень сходства с парадигмой индикатива (Табл. 3). Подпарадигма 2 л. полностью отлична от нее, формы 1 л. содержат показатель презенса на *-i*, но особые лично-числовые показатели; формы 3 л. — особый показатель, но те же лично-числовые: 2Р > 1Р, 3Р.

Табл. 3. Парадигма презенса индикатива, глагол ‘идти’

	SG	PL
1	ηən- <i>i</i> -i	ηən- <i>i</i> -pu
2	ηən-i-si	ηən-i-su
3	ηən-i-ni	ηən-i-ti

На первый взгляд небольшое, различие между формами системы, наблюдаемой в наших полевых данных, и системы Петровой / Суника ведет к куда более заметному различию на уровне структуры парадигмы (Табл. 2). Это различие и вопрос о том, можно ли считать его результатом (микро)диахронического сдвига¹ и если да, то как он происходил, будет рассмотрено подробнее.

С типологической точки зрения интересны следующие свойства ульчского императива:

– совпадение формы гортатива с формой презенса индикатива 2 л. мн. ч. (2);

¹ Об истории императивных форм на уровне тунгусо-маньчжурской семьи см. [Alonso de la Fuente 2012].

– противопоставление внутри форм 2 л. Т. И. Петрова [1936: 57] обозначает их как Praesens vs. Futurum imperativi, но применительно к нашим данным точнее говорить о противопоставлении по референциальному статусу ситуации. Форма на *-səri* (4) выступает как более узкая форма «нереферентного императива» (‘ситуация, не локализованная во времени’).

Примеры

- (1) *nu aldač-i-ta* *aj=məčilə*
 ну рассказывать-PRS-IMP.1SG этот=как.будто
 ‘Ну, расскажу-ка вот так!’. (oab)
- (2) *munži tatoči-ru,* *mun-ži gəsə*
 1PL-INS учиться-IMP 1PL-INS вместе
tatoč-i-su
 учиться-PRS-HORT
 ‘Учись с нами, давай вместе будем учиться!’. (aid)
- (3) *ujlə jaja ujsi-ži-ni* *xaj...*
 сверху песня звучать-JUSS-3SG это
 ‘Наверху песня пусть звучит!’. (zp)
- (4) *ulə muruču-sər*
 хорошо думать-IMP2
 ‘Хорошо думай!’. (lpd)

Литература

- Аврорин В. А. 1961. Грамматика нанайского языка. Т. 2. М. — Л.: Наука.
 ван дер Аувера Й., Гусев В. Ю., Добрушина Н. Р. 2004. Семантическая карта императива / гортатива // Типологические обоснования в грамматике / А. П. Володин (ред.). М.: Знак.
 Гусев В. Ю. 2013. Типология императива. М.: Языки славянской культуры.
 Петрова Т. И. 1936. Ульчский диалект нанайского языка. М.: Учпедгиз.
 Суник О. П. 1985. Ульчский язык: исследования и материалы. Л.: Наука.

- Alonso de la Fuente J. A. 2012. The first imperative of Tungusic // *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 129. P. 7–34.
- van der Auwera J., Dobrushina N., Goussev V. 2013. Imperative-Hortative systems // *The World Atlas of Language Structures Online* / M. S. Dryer, M. Haspelmath (eds.). Leipzig: Max Planck Institute for Evolutionary Anthropology.

Показатели каритива в лесном и тундровом энецком: сколько их?

А. Ю. Урманчиева

(ИЛИ РАН, Санкт-Петербург / ТГУ, Томск)

В лесном и тундровом энецком (< северносамодийские < самодийские < уральские) представлен ряд производных с каритивным значением (отсутствие некоторого предмета). В существующих описаниях лесного энецкого отмечаются следующие формы каритива: в [Сорокина 2010: 148–150] отмечена только адвербиальная форма *-śida/-đida/-tída* (где отмечается, что показатель является составным, и второй его элемент на синхронном уровне совпадает с одним из причастных показателей), в [Siegl 2013: 171–172] отмечено, что в лесном энецком существуют:

- во-первых, отыменные глаголы со значением необладания с показателем *-śi/-đi/-tí*:

enči *osa-śi*

person meat-CAR.3SG

‘The person is without meat’.

[Siegl 2013: 172],

- во-вторых, регулярная причастная форма от таких глаголов с показателем *-śida/-đida/-tída*, употребляющаяся в атрибутивной функции:

osa-ši-ďa

enči

meat-CAR-PTCP.IPF

person

‘person without meat’ (Lit. ‘meatless-being person’)

[Siegl 2013: 172],

– а также сказано, что эти две формы «глагольного» каритива «не надо смешивать с прототипическим именным каритивом с показателем *-šud/-ďud/-tūd*, который, как и следовало ожидать, синтаксически употребляется в обстоятельственной функции»:

tori-jet

bednye

kari-šud

such-EMPH

poor[ADJ.RU]

fish-CAR

kaji-bi-ši-jet

remain-PERF-3PL.PST-EMPH

‘So, the poor remained without fish’. [LDB Fishermen]

[Siegl 2013: 172].

Таким образом, в [Siegl 2013: 171–172] перечислены представленные в лесном энецком три формы каритива и описаны их синтаксические функции. Если взаимосвязь первых двух форм вполне очевидна и отмечена в обоих процитированных описаниях, отношение этих двух форм к форме с показателем, зафиксированном в работе Ф. Зигля как *-šud/-ďud/-tūd*, остаётся неясным. Является ли эта последняя форма морфологически элементарной? Связана ли она по происхождению с двумя первыми формами с каритивной семантикой?

Словарные материалы Е. А. Хелимского по тундровому энецкому языку из архива Института финно-угроведения Гамбургского университета позволяют уточнить формы каритивных показателей: за исключением разницы в форме причастного показателя (*-ďa* в лесном энецком, *-ďe* в тундровом энецком) они совпадают для этих двух языков. Вербальная репрезентация каритива имеет показатель *-se/-ďe/-tē*, производная от нее при помощи причастного показателя атрибутивная репрезентация имеет показатель *-seďe/-ďeďe/-tēďe*, ср. *ada-*

‘мысль’ > *adase-* ‘быть глупым’ > *adasede* ‘глупый’. Адвербиальная репрезентация имеет показатель *-šudiʔ/-đuʔ/-tuʔ*, ср. *sei* ‘глаз’ > *seišudiʔ* ‘вслепую’. В сравнении с материалами Ф. Зигля в моих материалах по лесному энецкому каритивные показатели зафиксированы в форме, более близкой к той, которая приведена Е. А. Хелимским для тундрового энецкого, ср. следующие записи от А. С. Пальчина: атрибутивная репрезентация *seiseda* ‘слепой’ (<*sei* ‘глаз’), *siʔorosedə* ‘немой’ (<*siʔoro* ‘язык’), адвербиальная репрезентация *tiešudiʔŋa* ‘он без оленей’.

Таким образом, архивные словарные материалы Е. А. Хелимского по тундровому энецкому и мои полевые словарные материалы по лесному энецкому позволяют уточнить фонетический облик трёх каритивных показателей энецкого языка. С опорой на это в докладе будут предложены прасамодийские праформы и этимология трёх энецких каритивных показателей. В частности, будет показано, что, несмотря на несходство на синхронном уровне, в диахронии все три показателя оказываются связаны между собой.

Литература

- Сорокина И. П. 2010. Энецкий язык. СПб: Наука.
 Siegl F. 2013. Materials on Forest Enets, an Indigenous Language of Northern Siberia. *Mémoires de la Société Finno-Ougrienne*, 267. Helsinki: Société Finno-Ougrienne.

Documentation of Negidal: Navigating through the data

N. Aralova¹, B. Pakendorf¹

(1 – CNRS / Université de Lyon / Kazan Federal University;

2 – CNRS / Université de Lyon, France)

Our talk has two goals: first, we would like to draw the attention of documentary linguists to the importance of organizing collected data in a transparent way that is accessible to their colleagues; secondly, we would like to present the collection of

Negidal materials which we have been archiving with the Endangered Languages Archive (ELAR) at SOAS, University of London (<https://elar.soas.ac.uk/>). We will pay special attention to our principles of data organization and metadata.

We consider both the language community as well as the wider academic community as important target audiences of documentation work. For linguists, data from linguistic documentation is the most precious source of knowledge about lesser known, underdescribed and endangered languages. As Seifart et al. [2018: e328–e329] put it, “the documentation of linguistic diversity keeps turning up new phenomena that had either been considered impossible or simply had not been contemplated as linguistic categories”. Needless to say, open access and a clear way of organizing the materials are crucial in making the documentation available to the academic community. Our recent experience of searching for data for a typological study shows that despite the growing number of documentation projects, quite lot of archived materials are hardly usable for one’s peers. Unfortunately, if the data are not archived properly all the efforts of the documentation are nullified.

In our documentation of Negidal, we are striving to archive the materials in such a way as to be accessible to linguists, to other academics, as well as to interested lay people. At time of writing (May 2019), we have archived about 11 hours of transcribed, translated, and glossed audio- and videomaterials [Pakendorf, Aralova 2017]. All the data are available upon registration with the ELAR archive. The design of ELAR allows only a flat structure of the the archived data, but we use the option of defining keywords, genres, and topics as a means to structure our deposit, as we will illustrate during our talk with concrete examples taken from our archive. This helps the user to get an overview of the contents of the corpus, which currently counts more than 130 recording sessions. The genres subdivide the data into larger categories, such as folklore materials or procedural texts, while the topics and

keywords enable users to find recordings dealing with particular subjects they are interested in, e.g. taboos and traditions or hunting.

A set of files related to one recording session is called a bundle. For each such bundle we provide metadata with information on the date of recording, the speaker(s) and the other participants involved in the work on this recording, as well as a very short summary of the contents. Since it is currently impossible to add keywords in Russian or switch the language of the ELAR interface to Russian, we are providing the metadata not only in English, but also in Russian, so that the Negidal community can use the archive, too.

We believe that the unique data archived within this project can indeed be of use for both the language community and the scientists interested in this Siberian people and their language.

Bibliography

Pakendorf B., Aralova N. 2017. Documentation of Negidal, a nearly extinct Northern Tungusic language of the Lower Amur. London: SOAS, Endangered Languages Archive.

URL: <https://elar.soas.ac.uk/Collection/MPI1041287> (accessed on: 29 July 2019).

Seifart F., Evans N., Hammarström H., Levinson S. C. 2018. Language documentation twenty-five years on // *Language* 94(4). P. e324–e345. URL: https://www.linguisticsociety.org/sites/default/files/e05_94.4Seifart_0.pdf

(Re-)Defining the legacy: integration of the Tomsk language archive data into modern LDD projects on endangered Siberian indigenous languages

A. Filchenko

(Nazarbayev University, Kazakhstan)

Siberian languages have a fair research tradition, pioneered as a purposeful systematic LDD activity by the likes of M. A. Castren in the 19th century. The 20th century saw gradual increase of academic

interest and the number of programs aimed at bridging the many empirical gaps concerning the languages of the region. Notably so was the series of projects undertaken in the 1950s–1970s by Tomsk Siberian Indigenous Languages Laboratory led by the prominent local linguist and archeologist Andreas Dulson. The program pursued a long-term interdisciplinary documentation of the aboriginal Siberian idioms of the so-called Ob-Yenisei linguistic area, which resulted in a unique collection, the Tomsk Field Archive of Siberian Indigenous Languages (totalling over 180 volumes of handwritten field-notes and hours of analogue audio-recordings on at least 8 aboriginal Siberian languages). Unique for its time and region, and collected within the contemporary state-of-the-art uniform theoretical and methodological framework, the archive was maintained and supplemented throughout 1980s–1990s as a working empirical database for a series of narrowly defined, primarily descriptive research projects, and a very limited set of language maintenance initiatives.

The value and interest to the Archive rose considerably in the early 2000s, conditioned by two important factors: the dramatic rate of local aboriginal language endangerment / extinction; and the advent of modern international endangered language documentation programs such as DOBES, ELDP, and NSF-DEL, among others. Beginning in 2005, the Archive featured in a number of documentation projects, offering an important empirical contribution to the debates on the evolution, variation and change in the languages of the region. While naturally focused on fieldwork in relevant indigenous communities and on collection of primary language data, these projects had as an important component the “re-archiving” (digitization and modern re-annotation) of legacy data on the respective languages.

The experience of these projects has been informative for the discussion of such issues as the life and the role of legacy data in modern LDD initiatives. The range of technical, methodological and ethical issues associated with integration of legacy data

archives into modern digital endangered language maintenance and revival projects has not yet been an object of special attention. The paper reviews the activities, problems, decisions, and processes of integrating legacy materials of the Tomsk field archive into modern LDD programs on endangered Siberian indigenous languages exemplified by two projects on Siberian Uralic and Turkic languages. Such questions will be explored and illustrated as: re-invention of the function and value of the legacy data in light of modern emphasis of empirical verifiability and open data access; application of legacy data to new research questions and methods; re-appropriation of legacy data in altered form and representation by academic community, and by the language communities for maintenance and revitalization.

Profiling Dolgan before its standardization

S. Fujishiro

(Kobe City College of Nursing, Japan)

The Dolgan language is a minority Turkic language spoken in Taymyr, in the far northern part of the Krasnoyarsk region, and in the Anabar area of Sakha-Yakutia, the Russian Federation. There were approximately 5000 speakers of Dolgan according to the 2002 census, but the number decreased significantly by the 2010 census to about 1000. This language evolved recently, in the 18th or 19th centuries, in Taymyr, where a number of Evenki, Yakut, and Nganasan people, as well as the so-called Russian “tundra-peasants,” formed a community and selected Yakut as their common language. The form of Yakut spoken in this emergent “Dolgan” community or ethnic group developed into the Dolgan language. Dolgan provides a model of the complicated ecological process of small languages in Siberia: language formation — standardizing — endangerment.

Linguistically, today’s standard Dolgan might be regarded as a dialect of Yakut; thus, it might also be called the Taymyr dialect

of the Yakut language. It does not differ very radically from standard Yakut, with the main differences in the lexicon and some minor differences in grammar and phonetics.

In this paper, I report on one of the characteristics of the language before Dolgan was standardized in the latter half of 20th century, based on the Dolgan-Russian Card Dictionary (hereinafter DRDC) that was compiled by K. M. Rychkov in the first decade of the 20th century (unpublished, preserved in the archive of Institute of Oriental Manuscripts, Russian Academy of Sciences) and on a Yakut Dictionary (hereinafter, DYL) that was edited by E. K. Pekarskij and published between 1907–1930. These two dictionaries mostly contain words collected in the latter half of the 19th century and the first decades of the 20th century.

The DRCD has 1534 entries (cards) that were collected by Rychkov during his expeditions; the dictionary was completed in 1911. The DYL, edited by Pekarskij, has about 25000 entries and was compiled in the latter half of the 19th and the first half of the 20th centuries based on data collected by various researchers, activists, and others, as well as by Pekarskij himself. Among these, there are about 300 entries, words and expressions that, as we can infer, were introduced into the DYL from Dolgan mainly from Tajmyr. For example, some of them are annotated with notes such as *у долгань* ‘(spoken) by Dolgans’; some others are marked with (*Вас.*), which means that the word or expression was found among the descriptions by a researcher V. N. Vashiliev (Васильев) in his expedition to the Dolgan-Yakut people in the Enisej region in 1905; and so on.

It is difficult to identify and define the status of a language, especially when we do not have enough linguistic data, as with the languages of Imperial Russia in Siberia. In this paper, we examine these 300 entries of the DYL along with the 1534 entries of the DRCD, keeping in mind that the 300 entries somehow deviated from “ordinary” Yakut entries from the time. The results will contribute to profiling Dolgan prior to its standardization and

providing some support to the idea that Dolgan could be a model of the complicated ecological process of small languages in Siberia.

Bibliography

- Пекарский Э. К. 1958 [1907–1930]. Словарь якутского языка. АН СССР.
Убрятова Е. И. 1985. Язык норильских долган. Новосибирск: Наука.
Fujishiro S. 2011. The process of formation of the Dolgan ethnicity and language (Based on material of K. M. Rychkov from Enisej Region of the former Russian Empire (I)) [in Japanese] // CSEL series. Vol. 17. Kobe. P. 175–214.

Variation in two dictionaries of Botlikh

G. Moroz, Ch. Naccarato, S. Verhees
(*NRU HSE, Moscow*)

Botlikh is an unwritten language of the (Avar-)Andic branch of East Caucasian, spoken mainly in three villages in western Daghestan (Botlikh, Miarso and Ashino). Estimations of the number of speakers vary from 4000 [Alekseyev 2016: 203] to 8000 [Saidova, Abusov 2012: 17]. Remarkably, two dictionaries are available for the language.

Both dictionaries were compiled independently in the 2000s and contain 8000–9000 head words and a wealth of examples. The materials in the dictionary by Saidova and Abusov [2012] were collected by Magomed Abusov (a native of Botlikh), who consulted with various speakers in the period 2003–2009. The dictionary of Alekseev & Azaev [to appear] is based on the card library created by Khalil Azaev (another Botlikh native), who wrote the first major study on the Botlikh lexicon [Azaev 1975]. It is unclear how many speakers were consulted exactly, and what their respective ages and backgrounds were.

In our talk we discuss how we merged the two resources, and how this information can be used to study language variation and change of an otherwise understudied language. We studied the

frequency of phonological units (segments, syllables, stress patterns), and compared information on the inflectional paradigms of nouns and verbs (formation of oblique stems and plural forms, basic tense forms). The dictionaries revealed considerable variation in both areas. In some cases it concerns free variation and possibly speakers' idiolects, whereas other differences seem to reflect diachronic changes (the material collected by Azaev is slightly older).

Bibliography

- Alekseev M. E., Azaev X. G. (to appear). Botlixsko-russkij slovar [Botlikh-Russian Dictionary]. Moscow: Academia.
- Alekseyev M. 2016. Botlikh // Word-Formation. An International Handbook of the Languages of Europe. Vol. 5 / P. O. Müller, I. Ohnheiser, S. Olsen, F. Rainer (eds.). Berlin — N.Y.: De Gruyter Mouton. P. 3678–3685.
- Azaev X. G. 1975. Leksika i slovoobrazovanie botlixskogo jazyka [Lexicon and Word-formation in Botlikh]. Ph.D. dissertation. Tbilisi State University.
- Saidova P. A., Abusov M. G. 2012. Botlixsko-russkij slovar' [Botlikh-Russian Dictionary]. Makhachkala: Institute of Language, Literature and Arts of Dagestan Scientific Centre, RAS.

Spatial cases in Udihe: a quantitative analysis

E. Perekhvalskaya
(*ILS RAS, Saint Petersburg*)

In Udihe, the case system includes several spatial cases which express the semantic category of orientation. There is some uncertainty in the precise definitions of the case meanings, mainly in making a demarcation line between the usages of Locative and Directive, on the one hand, and between Locative and Dative, on the other. The fundamental Udihe grammar says: “The distribution of the Locative and the Dative in the local sense does not follow any strict semantic or phonological criteria and has to be learnt for

each individual instance” [Nikolaeva, Tolskaya: 2004, 125]. Nikolaeva and Tolskaya suggest that the choice between the Locative and the Dative depends on a particular NP. They note that the nouns *kniga* ‘book’ and *zugdi* ‘house’ usually require the Dative: *kniga-du* ‘in the book’ and *zugdi-du* ‘in the house’, while nouns *uli* ‘water, river’ and *namu* ‘sea’ are usually used in the Locative form: *uli-lə* ‘in the river’, *namu-la* ‘in the sea’.

Functions of the cases as well as the verb control are rather difficult to study using modern data. It should be noted that the method of elicitation (translation of separate sentences) is not suitable. All modern native Udihe speakers have Russian as the dominant language. Therefore, when translating from Russian into Udihe, they model Russian constructions. For example, in Udihe the distribution of actants with verbs of speech is as follows: the speaker is coded by Nominative, the reported information is a direct object (Accusative), the addressee is the indirect addition (Directive case):

- (1) *Ain-tigi-hi* *mine-wə* *əʒi* *təluʒusi*.
 brother-DIR-2SG 1SG-ACC VNEG.IMP tell
 ‘Don’t tell your brother about me!’.

Modern speakers use the Dative case to encode the addressee: ‘tell the brother’ is translated as *aga-du dianami*, following the Russian model *skazat’ brat-u* (DAT). The form *aga-du* is also impossible in traditional (“old”) Udihe because the appellative term *aga* ‘brother!’ is used as a referential term.

In this situation, the main method of study is corpus analysis.

I have analyzed the uses of the three cases (Dative, Locative and Directive) in the corpus of texts (40000 words). Several nouns were chosen: *bua* ‘place, nature’, *zugdi* ‘house’, *xokto* ‘road’, *uli* ‘river, water’, *kää* ‘edge, bank’.

The usage of caseforms of these noun with different verbs was analysed. The analysis showed that the use of cases does not depend on the particular NP. Consider the following example:

ʒugdi ‘house’:

Loc *ʒugdilā* — 99 occurrences, with verbs: *ii-*, *iigi-* ‘to reach’ — 59; *ḡənə-*, *ḡəni-* ‘to go (back)’ — 19; *əmə-* ‘to come (back)’ — 10; other verbs of motion — 3.

In 92% of occurrences *ʒugdilā* is used with verbs of motion (60% with the verb *ii-*, *iigi-* ‘to reach’). Among the rest are verbs *ḡua-* ‘to sleep’ (2) and *aḡasi-* ‘to spend night’ but the both are used in Imperative with the meaning ‘come and sleep’, ‘come and spend the night’.

Dat *ʒugdidu* — 11 occurrences, with verbs: *bii-* ‘to be, to live’ — 4; *bagdi-* ‘to live’ — 2; no verbal predicate — 2; *əkə-* ‘to stay’ — 1; *inə-* ‘to spend the day’ — 1; *umi-* ‘to drink’ — 1.

About 90% of occurrences of the Dative are with verbs of position or existence.

The same calculations were made for the case forms of other chosen nouns.

The analysis showed that Dative expresses the meaning of quiescent state and most often matches with the verb *bii-* ‘to be, to live’. Locative expresses the attained aim of motion. This aim may be perceived as “position”, but unlike an NP in Dative, it designates a position reached as a result of motion. Note that Locative does not point to the precise localization of the action, e.g. *ḡyḡə-lo-ni saṇzehæwa tuləsiiti* ‘they used to put a nasal ring **into his nose**’; *Bəli-lə ihigehæti* ‘they reached **Khabarovsk**’.

The analysis of the Locative and Directive cases showed that the main distinction between the Locative and the Directive is whether the reference point or the goal is achieved, e.g.:

Dative: *ʒugdi-du bihi nii budəiti* ‘people who were **at home** were dying’

Locative: *ʒaṅgæ ʒugdi-lə-ni iigihæti* ‘the chief <and his people> entered **the house**’;

Directive: *ʒugdi-tigi-i ḡenihəni* ‘he set off **for his house**’.

Thus, the choice between the three spatial cases in Udihe is not random; it depends on the linguistic interpretation of the situation in question.

Possession in Ob-Yenisei Languages: documentation-to-typology-to-documentation

O. Potanina, A. Filchenko

(Nazarbayev University, Kazakhstan)

A set of documentation projects on North-Western Siberian indigenous languages comprising the so called Ob-Yenisei linguistic area has led to a subproject on local typology of possession. These lesser described and highly endangered idioms do feature as datapoints in the large typological databases such as WALIS, however, for various reasons, these datapoints may not necessarily represent the actual diversity of these systems adequately, calling for a more representative local typology study. Such local typology project focusing on possession has been developed based on three conceptual statements: 1) there are no comprehensive, theoretically and methodologically consistent typological descriptions of possession in the Ob-Yenisei languages that take into consideration areal factors and language contact; 2) there are no exhaustive typological parameters defined for the cross-linguistic study of possession; 3) there are unpublished language documentation data on these highly endangered dialects, largely unavailable to academic community.

The languages of the project include Uralic (Eastern Khanty, Central Selkup, Nganasan), Yeniseian (Central Ket), Turkic (Chulym Turkic, Bacht Teleut).

On the one hand, by creating a “typological portrait” of possession in Ob-Yenisei languages, the project conducts a theoretically and methodologically uniform description of a wide range of possessive constructions, in structurally diverse, genetically related and unrelated languages of the area. On the other

hand, the project seeks to design a typological questionnaire for the study of possession (akin to MPI Leipzig questionnaires), thus concerned not only with mere illustration a wide range of morphosyntactic strategies employed by these languages, but also with describing and interpreting the full diversity of possessive constructions in terms of general typological perspective, grammaticalization theory, information structure theory, utilizing typological parameters proposed for the description of adnominal and predicative possession by M. Koptjevskaja-Tamm, L. Stassen and B. Heine, among others.

Being a universal category, possession does not have universal ways of manifesting possessive relations: lexical, morphological, syntactic. The notion of possession remains ambiguous. Besides formal variation, possessive constructions have various functions and meanings: ownership, part-whole, kinship, associative possession, and pragmatic functions of definiteness / indefiniteness, specificity, salience.

The project is still under way and is currently reviewing the intra- and extra-linguistic principles that condition the variety of possessive constructions in the languages, establishing the parameters valid for cross-linguistic comparison. Within this variety of possessive constructions, the project identifies basic patterns according to the following criteria:

- if the construction is restricted to a grammatical subtype, it is less basic;
- if the construction is semantically or pragmatically specialized, then it represents a less basic type;
- more complex constructions represent a less basic type;
- less frequent is less basic.

Thus, among the formal parameters for adnominal possessive constructions valid for cross-linguistic comparison, the project identifies:

- i) the order of the possessor and the possessed;
- ii) the locus of marking of possessive relations;
- iii) means of expressing the possessor and the possessed;

iv) means of coding the possessive relations (degree of explicitness, degree of fusion).

The values of these parameters are conditioned by the following principles:

- if the possessor is: human or non-human, preferential or non-referential, pragmatically salient or inactive in discourse universe.
- if the possessee is: alienable or inalienable, definite or indefinite.

On the relationship of documentation, dialectology, and standardization: materials from Tundra Nenets

T. Salminen

(University of Helsinki / Kone Foundation, Finland)

It seems safe to say that for the majority of languages spoken in the world, documentation of their dialects has been carried out unevenly. Explanatory factors behind this unevenness are varied, but concern the research history of the language in general, the role of the standard variety of the language, and the extent of dialectal differentiation within the language. The Tundra Nenets language is no exception to these issues.

The early documentation of languages without old literary traditions was by necessity essentially dialectological. For practical reasons, the dialects spoken near colonial outposts and in other easily accessible locations tend to be most thoroughly studied, barring individual efforts of pioneer linguists to expand the geographic range of their activities. The example provided by the history of fieldwork on Tundra Nenets makes it clear that the dialects spoken close to Pustozersk (Naryan-Mar) and Obdorsk (Salekhard) appear among the best known, but also that the 19th century and the early 20th century explorers were able to reach some more remote areas as well. Large parts of the Tundra Nenets country, however, remained unexplored until much later, and efforts to study individual dialects have continued till the present day, most recently in the

form of two remarkable dialectological dictionaries dedicated to the Kanin dialect [Бармич 2015] and the Yamal dialects [Буркова и др. 2010], respectively. All in all, through an analysis of multiple and diverse sources, it is possible to capture a reasonably detailed description of Tundra Nenets dialectal variation.

It seems widely assumed that standard varieties are influencing and to an extent eradicating non-standard dialects across the globe. A more accurate claim might be that this happens when the standard is based on the dialect of the capital or other core region, while mixed or artificial standards, not immediately associated with a particular spoken prestige variety, generally lack the potential to level out dialect differences. A standard form of Tundra Nenets may be said to have been crystallized in the famous “red dictionary” [Терещенко 1965], but, since it recognizes a fair amount of variation, typically what is found in Bol’shaya Zemlya vs. Yamal dialects, the Tundra Nenets literary language may well be characterized as pluricentric, and few signs of dialect levelling have been discovered. At the same time, however, there are indications that in a number of field projects the Tundra Nenets consultants employed forms they considered standard instead of relying on their native competence, as will be discussed in detail in the presentation.

The number of dialectal differences within Tundra Nenets is relatively high, yet the vast majority of them appear straightforward and transparent. Besides discussing general aspects of the topic, the presentation contains a concise description of the distribution and origin of Tundra Nenets dialectal variation, with a focus on features that remain poorly understood.

Bibliography

- Бармич М. Я. 2015. Неян вабц’ вада” — Словарь маминого языка. СПб: Алмаз-Граф.
- Буркова С. И., Кошкарева Н. Б., Лаптандер Р. И., Янгасова Н. М. 2010. Дialectологический словарь ненецкого языка. Екатеринбург: Баско.

Терещенко Н. М. 1965. Ненецко-русский словарь. М.: Советская Энциклопедия.

University of Tartu archives of Estonian dialects and kindred languages: materials of Uralic languages spoken in Russia

T. Tuisk

(University of Tartu, Estonia)

This paper introduces the materials of Uralic languages spoken in Russia found at the Archives of Estonian Dialects and Kindred Languages (AEDKL). AEDKL is a collection of Uralic linguistic materials located at the University of Tartu that consist of fieldwork recordings, written materials, photos and videos of Estonian Dialects, Finno-Ugric and Uralic languages [Lindström, Lippus, Tuisk (to appear); Rätsep 2003].

The archives contain four types of materials: **1)** sound recordings (Finnic: Estonian, Livonian, Votic, Ingrian, Veps, Karelian, Finnish, Ingrian Finnish; Finno-Ugric: Inari Saami, Erzya, Moksha, Komi, Udmurt, Hungarian, Khanty; Samoyedic: Nenets, Kamas); **2)** unpublished manuscripts, including student papers and thesis defended at the Institute of Estonian and General Linguistics, fieldwork diaries, transcriptions and written notes on the Uralic languages; **3)** photos from fieldwork expeditions and linguistic events; **4)** video recordings. All materials in AEDKL include sound recordings, manuscripts, photos and videos from the Uralic languages spoken in Russia.

The archives consist of about 2800 hours of fieldwork recordings of the Uralic languages. The majority of the sound recordings are of the Estonian dialects with the remainder composed of recordings of other Finno-Ugric languages. The earliest sound recordings date back to the 1950s and written materials, to the 1920s. There are a total of 392000 pages of written manuscripts in the archives (~ 266000 pages are digitally available). The collection

holds about 2900 photos from fieldwork expeditions and linguistic events (e.g., conferences and seminars). Photos are divided into two series based on media type: paper and digital photos. There are around 1300 paper photos that are digitized, and digitization is still in progress. Around 1600 digital photos are from recent years of fieldwork and different linguistic events. Video recordings are from fieldwork conducted during recent years. Also, old film rolls from the 1970s and 1980s have been digitized (these include recordings made during fieldwork and at various university events). There are 51 hours of video recordings, but in the current version of the database viewing these videos is not integrated into the online archive system.

The organization of the collection began in 2000 with the digitization of the sound recordings. The work moved on to scanning the written materials and photos, and organizing the metadata into an online database. The online database of the archives is freely accessible and open to all researchers at www.murre.ut.ee/arhiiv/. By now, the majority of the materials are digitized, but the work on the archives and the database continues.

Bibliography

- Lindström L., Lippus P., Tuisk T. (to appear). The online database of the University of Tartu Archives of Estonian Dialects and Kindred Languages and the Corpus of Estonian Dialects // *Uralica Helsingiensia* series. Plurilingual Finnic. Change of Finnic Languages in a Multilingual Environment / S. Björklöf, S. Jantunen (eds.). Helsinki: Suomalais-Ugrilainen Seura [Finno-Ugrian Society].
- Rätsep H. 2003. Tartu Ülikooli eesti keele arhiivi saamisloost ja saatusest [On the origin and destiny of the Estonian language archive at the University of Tartu] // 200 aastat eesti keele ülikooliõpet. Juubelikogumik. Tartu Ülikooli eesti keele õppetooli toimetised. 25. P. 153–170.