

Научный центр по сохранению, возрождению и документации языков России

Институт языкознания РАН

**Итоговые документы, подготовленные в ходе научно-исследовательской
работы
“Разработка плана и методик сохранения и возрождения языков России”**

**7. Оценка документации языков России, находящихся под угрозой
исчезновения**

Н. Б. Аралова, Е. М. Будянская, А. А. Сюрюн

Оглавление

1. Введение
2. Задача исследования и смежные вопросы
 - 2.1. Формулировка задачи
 - 2.2. Актуальность документации исчезающих языков
 - 2.3. Параметры анализа
 - 2.3.1. Количество носителей
 - 2.3.2. Имеющаяся документация
 - 2.3.3. Существование исследовательских групп
 - 2.3.4. Существование активистов, работа по поддержке языка, запрос от языкового сообщества на сотрудничество с лингвистами
 - 2.4. Общая формулировка критерия для выбора языков
 - 2.5. Целесообразность и актуальность поставленной задачи
3. Ход работы
 - 3.1. Подготовительный этап
 - 3.2. Разработка параметров
 - 3.3. Сбор информации
4. Языки или диалекты, для которых наиболее актуальна документация в 2022 году
 - 4.1. Языки, для которых актуальна собственно полевая документация
 - 4.2. Языки, для которых актуально социолингвистическое исследование
 - 4.3. Языки, документация которых проводится в 2021 г. на базе Научного центра (пилотный проект)
5. Другие языки или диалекты в нашем обзоре
 - 5.1. Языки, для которых актуально социолингвистическое исследование
 - 5.2. Языки, для которых актуальна работа с имеющимися данными
 - 5.3. Языки, для которых актуальна работа с имеющимися материалами и их пополнение
 - 5.4. Языки, для которых актуальна собственно документация
6. Заключение
7. Благодарности экспертам
8. Литература

Приложение 1. [Обзор документации по исчезающим языкам России.](#)

Приложение 2. [Н. Б. Аралова, Е. М. Будянская, А. А. Сюрюн, Е. Ю. Груздева. Как документировать языки, чтобы их сохранить. Методические рекомендации по сбору материала для потенциальной поддержки и ревитализации миноритарных языков.](#)

1. Введение

Настоящий документ подготовлен группой исследователей, работающих в Центре по сохранению, возрождению и документации языков России при Институте языкознания РАН, перед которыми была поставлена задача оценить задокументированность языков России, находящихся под угрозой исчезновения, составить список языков, требующих срочной документации, а также выработать рекомендации для остальных языков. Данная задача выполнялась в рамках подготовки Программы по сохранению и возрождению языков народов России, разрабатываемой в Институте языкознания РАН.

В ходе работы был проведен подробный обзор документации исчезающих языков России. Исследование включало уточнение релевантной социолингвистической информации по каждому из рассматриваемых языков, выработку критериев по оценке документации (наличие словарей, грамматик, периодики, художественной литературы, коллекций текстов, научных групп), сбор информации по доступным источникам, работу с экспертами по конкретным языкам.

В результате исследования были проанализированы данные по 48 языкам и составлен список из десяти идиомов, требующих срочной документации в 2022 г. (ительменский, южноюкагирский, науканский, хамниганский, сойотский и южнорутульский (борчинско-хновский) языки; амухский и анклухский диалекты сирхяцудахарского языка, аганский говор нешанского языка и северносахалинский диалект амурского нивхского языка).

В итоговом документе, к которому прилагается аналитическая таблица (Приложение 1), рассматривается общая проблематика документации, обсуждаются критерии ее оценки, приводятся обоснования для отбора языков для собственно полевой и социолингвистической документации, а также даются рекомендации по документации идиомов, не вошедших в первоначальный список.

Кроме того, исследовательская группа подготовила общие методические рекомендации по сбору материала для потенциальной поддержки и ревитализации миноритарных языков, которые были опробованы в летних экспедициях коллег-лингвистов (Приложение 2).

В качестве практической апробации предлагаемых методик осуществляется пилотный проект по документации уильтинского (орокского) языка, который находится на грани исчезновения. Документ включает информацию о реализации этого проекта.

2. Задача исследования и смежные вопросы

2.1. Формулировка задачи

Основной задачей нашей рабочей группы на данном этапе стал обзор документации языков России, которые находятся на грани исчезновения. К настоящему моменту в нашей стране накопился большой опыт по языковой документации, существует ряд научных групп и центров, которые занимаются различными языками в сложной социолингвистической ситуации: это и сотрудники Института языкознания РАН, и сотрудники и учащиеся других научных и образовательных учреждений Москвы (экспедиции кафедры теоретической и прикладной лингвистики МГУ, Школы лингвистики НИУ ВШЭ, Института лингвистики РГГУ; см., например, проекты Международной лаборатории языковой конвергенции и Лаборатории социогуманитарных исследований Севера и Арктики НИУ ВШЭ) и других городов (Санкт-Петербург, Томск, Новосибирск и др.). Поэтому на первом этапе работы нашей группы наиболее актуальной задачей стал анализ ситуации с документацией малых языков: нам было необходимо определить, насколько задокументированы языки России, находящиеся на грани исчезновения, есть ли среди них такие, которые необходимо срочно документировать, а активных проектов по документации не ведется. На основе этого анализа мы выбрали потенциальные направления для работы начиная с 2022 года.

Параллельно с этим мы разрабатываем практические рекомендации по документации малочисленных языков (см. Приложение 2) с особым вниманием на жанры, которые могут быть полезны для потенциальной ревитализации или же для создания учебных материалов.

Рекомендации основаны на литературе по документации, отражающей общемировой опыт, нашем собственном опыте и опыте пилотного проекта.

2.2. Актуальность документации исчезающих языков

Прежде чем перейти к обсуждению параметров анализа, необходимо сказать несколько общих слов о языковой документации в целом. После ежегодной встречи Американского лингвистического общества в 1991 году и статьи о проблеме исчезающих языков, вышедшей в журнале *Language* в 1992 году по итогам этого мероприятия (см. [Hale et al. 1992]¹), внимание лингвистов к активному процессу

¹ Эта статья представляет собой серию эссе разных авторов, объединенных общей тематикой.

исчезновения языков стало очень пристальным. В этой статье Майкл Краусс формулирует программу действий: собрать информацию о состоянии языков, заняться активной документацией тех языков, что находятся в самом критическом состоянии (и в первую очередь генетически или типологически изолированных), создать сеть архивов. К стандартному набору описания языка (грамматика, словарь, тексты) добавляется также возможность и необходимость аудио- и видеофиксации речи. Утверждается, что собранные в ходе документации материалы не только ценны с узко научной точки зрения, но могут сыграть важную роль и в ревитализации языка. Для языков, где еще не прекращена передача детям в семье, предполагается не только их документация, но и активное участие в поддержке и сохранении языка.

За прошедшие 30 лет была проведена огромная работа по документации языков мира, см., в частности, обзоры в [Austin 2016; Seifart et al. 2019]. Тем не менее, эта задача актуальна и сегодня.

Задачи, сформулированные в [Hale et al. 1992], можно представить как основные составляющие документации:

- Социолингвистическое исследование

Такое исследование позволяет понять, сколько человек владеет интересующим нас языком, на каком уровне, где они проживают, какого они возраста; дает возможность представить себе в целом и в деталях языковую ситуацию в исследуемой местности. Все это необходимо для планирования документации и других (социо)лингвистических и этнографических исследований.

- Собственно документация

Эта работа представляет собой непосредственный сбор материалов в ходе работы с носителями. Такую работу можно проводить в экспедиции или приглашая носителей к себе, или, что стало наиболее актуально в последнее время, делать удаленно.

- Обработка собранных материалов для того, чтобы ими могли пользоваться:
 - лингвисты и другие исследователи,
 - представители языкового сообщества.

Обработка материалов и предоставление открытого доступа к ним – очень важный этап документации. Без него сбор материалов в некоторой степени теряет смысл.

2.3. Параметры анализа

Чтобы оценить социолингвистическое состояние языков России и имеющуюся по ним документацию, мы выбрали ряд параметров, по которым мы постарались описать каждый из выбранных идиомов:

- 1) Количество носителей.
- 2) Имеющаяся документация и ее объем:
 - а) грамматики;
 - б) словари;
 - в) коллекции текстов и корпус.
- 3) Существование исследовательских групп, занимающихся документацией этого языка.
- 4) Существование активистов, ведущих работу по поддержке языка, и наличие запроса от языкового сообщества на сотрудничество с лингвистами.

Отдельно необходимо отметить, что информация, собранная нами в таблице “Обзор документации по исчезающим языкам России” (см. Приложение 1) по каждому языку, не является исчерпывающей. Нашей первоочередной целью было собрать сведения по каждому из указанных выше параметров, чтобы оценить общее состояние документации языков, находящихся под угрозой исчезновения. Ниже мы рассмотрим более детально каждый из перечисленных нами параметров.

2.3.1. Количество носителей

Несмотря на кажущуюся простоту этого параметра, дать точную оценку количества человек, владеющих тем или иным языком, бывает очень непросто. В первую очередь, это может быть связано с отсутствием информации о том или ином языковом сообществе (если социолингвистических исследований не проводилось, как в случае с хамниганами России). Кроме того, с точки зрения потенциальной документации важными становятся также такие факторы, как уровень владения языком, возраст, возможность и желание работать с лингвистами. Среди рассмотренных идиомов мы обнаружили языки, по которым имеется крайне мало собранных материалов, – калмакский язык и язык тазов. Однако уровень владения этими языками у последних носителей, к сожалению, уже не позволяет проводить с ними полноценную лингвистическую документацию. Возраст и здоровье потенциальных информантов также часто являются сдерживающим фактором для проведения документационных

проектов. Так, насколько нам известно, последние носители медновского алеутского уже не обладают достаточным здоровьем для активной работы с лингвистами (к счастью, Е. В. Головкин успел сделать с ними записи – сейчас он занимается обработкой этих материалов). Таким образом, оценка общего количества владеющих языком и возможностей работы с компетентными носителями того или иного языка является нетривиальной задачей. В условиях языкового сдвига количество носителей, свободно владеющих языком, обычно отличается в меньшую сторону от общего числа говорящих на языке. Существование нескольких носителей языка не обязательно гарантирует возможность работы по документации в силу причин, затронутых нами выше. Поэтому мы включили в нашу обзорную таблицу столбец, содержащий информацию о максимальном уровне владения языком и возрасте младшего носителя. Для языков, по которым удалось оценить состояние по диалектам, мы также указываем общее число носителей языка (как правило, по Всероссийской переписи населения 2010 года) и оценку количества владеющих данным диалектом.

2.3.2. Имеющаяся документация

Традиционно принято считать, что основные источники информации по языку – это грамматическое описание, словарь и коллекция текстов. Мы также разделяем это представление, однако обращаем внимание на более тонкие различия в данных категориях.

а) В блоке грамматического описания мы отмечаем, есть ли по данному языку академическая грамматика, учебная грамматика, грамматический очерк и работы на отдельные грамматические темы. В библиографию также включались значимые диссертации.

б) В словарном блоке мы отдельно отмечаем наличие академического, учебного и озвученного словарей. Если по языку имеется онлайн-словарь, мы добавляли и эту информацию тоже.

в) В блоке, относящемся к коллекции текстов, мы пытаемся оценить количество текстовых материалов на языке, при этом мы делаем различия между опубликованными печатными материалами, аудио- и видеоматериалами. Кроме того, параллельно мы стараемся оценить степень обработки этого материала (есть ли транскрипция и перевод, есть ли глоссы, представлены ли тексты в форме корпуса с возможностью поиска). Другой важный параметр – это доступность материалов: есть ли они в сети Интернет,

легко ли получить к ним доступ, насколько понятно оформлены метаданные и можно ли осуществлять по ним поиск.

Одним из самых сложных вопросов стало для нас определение “достаточности” документации. Интуитивно понятно, что окончательного предела в процессе документации не существует: дополнять материалы (как текстовые, так и словарные и грамматические) можно записями разных жанров, полученных от носителей разных полов и разных возрастных групп, представляющих разные диалекты и говоры, пока имеется такая возможность. Современные стандарты в области лингвистической документации предполагают наличие как аудио-, так и видеозаписей, а также озвученного словаря. В идеале каждый текст в коллекции должен сопровождаться морфологической аннотацией (глоссами) и переводом. Но при этом оценка полноты словаря и количества текстов на разных языках на данный момент скорее субъективна. Не совсем понятно, сколько часов звучащей речи или лексем в озвученном словаре является достаточной документацией, и в каких единицах эти параметры можно измерить. Мы не всегда можем адекватно оценить, достаточно ли подробно и исчерпывающе то или иное грамматическое описание или насколько полно собрана коллекция текстов. Эта оценка, видимо, должна производиться специалистами по соответствующим языкам, но и в этом случае она не будет лишена субъективности.

Отдельно следует подчеркнуть специфику проведения документации в рамках работы Научного центра по сохранению, возрождению и документации языков России. Одной из наших задач является документация, нацеленная не только на нужды академического сообщества, но и на сохранение языка в языковом сообществе, на совместную работу с языковыми активистами (ср. [Amery 2009; Austin, Sallabank 2018]). Мы стремимся к тому, чтобы материалы лингвистической документации могли быть использованы в пособиях по языку, в обучающих программах и были доступны языковому сообществу. Так, для большинства языков, включенных в обзор, отсутствует целый ряд жанров, необходимых и востребованных с точки зрения ревитализации и поддержки языков, — бытовая коммуникация, речевые формулы и речь, обращенная к ребенку, и др. Изучение и запись подобных материалов следует проводить даже в тех случаях, когда, как кажется, уже проведена существенная документация языка.

2.3.3. Существование исследовательских групп

Занимаются ли языком (или его диалектом) специалисты в настоящее время – еще один важный параметр для рекомендации активных действий по документации языка.

Мы постарались оценить существующую документационную активность по каждому из рассмотренных языков и диалектов. Если на данный момент нет активных документационных проектов и осталось очень мало свободно владеющих языком носителей, мы рекомендуем этот язык или диалект для дальнейшей документации, даже если ранее уже было собрано достаточно много материалов.

2.3.4. Существование активистов, работа по поддержке языка, запрос от языкового сообщества на сотрудничество с лингвистами

Этот параметр мы также считаем существенным. Если в языковом сообществе есть желающие заниматься своим языком, работать с лингвистами, перенимать практики документации и развивать документацию своими силами, для нас это является важным аргументом в пользу документации этого языка.

2.4. Общая формулировка критерия для выбора языков

Основным критерием для определения языка как наиболее нуждающегося в срочной документации было совпадение трех обстоятельств: 1) малое количество носителей со свободным уровнем владения языком (как результат прерванной межпоколенческой передачи языка); 2) слабая лингвистическая документация языка или ее отсутствие; 3) отсутствие активной работы по документации этого языка в настоящее время.

Несмотря на то, что в основе нашего выбора языков, рекомендованных для активной документации (см. Раздел 4), лежал именно этот трехсторонний критерий, в некоторых случаях мы принимали во внимание и другие факторы. Например, в случае ительменского дополнительный фактор заинтересованности сообщества играл более важную роль, чем наличие / отсутствие документации. В случае аганского говора нешанского языка, по которому ведутся исследования и имеется художественная литература, наиболее важным обстоятельством нам показалось уязвимое состояние этого говора по сравнению с другими говорами языка, а также отсутствие видеоматериалов в публичном доступе. Некоторые языки Дагестана были включены в обзор, поскольку они были рекомендованы экспертами для документации как находящиеся в неблагоприятной социолингвистической ситуации (хотя в этом регионе многие другие идиомы недостаточно задокументированы, в то же время их социолингвистическая ситуация требует уточнений). Таким образом, несмотря на,

казалось бы, строгий критерий и четко сформулированные параметры (см. 2.3), окончательное решение по рекомендациям о документации конкретных идиомов было принято не полностью объективно: и мы, и специалисты, с которыми мы консультировались, иногда субъективно оцениваем важность тех или иных факторов.

2.5. Целесообразность и актуальность поставленной задачи

Сбор информации о степени задокументированности языков не является какой-то исключительно новой деятельностью. В настоящее время имеется большое количество фондов, программ, а также интернет-порталов, где можно найти базовую информацию о языке, данные о социолингвистической ситуации, различные материалы по языкам. В частности, одним из первых информационных ресурсов о языках России является портал “Языки народов Сибири, находящиеся под угрозой исчезновения”², созданный сотрудниками отдела Севера и Сибири Института этнологии и антропологии РАН в результате прошедшего в 2005 году Круглого стола по проблемам языков коренных народов Сибири, находящихся под угрозой исчезновения. Сайт содержит информацию о 28 языках Сибири, находящихся на грани исчезновения, статистические данные по материалам переписей 1926–2002 годов, а также включает основную библиографию по указанным языкам. К аналогичным работам относятся проекты Лаборатории исследования и сохранения малых языков Института языкознания Российской академии наук под руководством О. А. Казакевич, образованной в 2019 году. На сайте “Малые языки России”³ содержатся материалы о структуре и функционировании малых языков России и их локальных вариантов.

На странице Фонда сохранения и изучения родных языков народов РФ, учрежденного указом Президента РФ в 2018 году, представлена интерактивная карта родных языков народов РФ⁴, а также дается самая общая социолингвистическая информация, справка об основных источниках (грамматики, словари, сборники или электронные корпуса текстов) по 138 языкам, распространенным на территории России.

Полезным для документации ресурсом является OLAC (Open Language Archives Community)⁵ — объединение открытых языковых архивов с целью создания унифицированного средства поиска информации по языкам мира в базах данных

² Портал доступен по следующей ссылке: <https://lingsib.iea.ras.ru/ru/>.

³ Сайт “Малые языки России” доступен по следующей ссылке: <https://minlang.site>.

⁴ Карта доступна по следующей ссылке: https://xn--d1acgejfp6hc6b.xn--p1ai/?page_id=765.

⁵ Ресурс OLAC доступен по следующей ссылке: <http://www.language-archives.org>.

различных учреждений и частных лиц. Данный цифровой архив представляет собой виртуальную библиотеку, в которой перечислены все доступные материалы по языку (список имеющейся литературы, ссылки на аудио и видео).

Еще одним проектом, аккумулирующим информацию по языкам, находящимся под угрозой исчезновения, является Endangered Language Project (ELP)⁶. Информация о языках для этого проекта берется из The Catalogue of Endangered Languages (ELCat) — каталога, работа над которым ведется Гавайским университетом в Маноа и Университетом Восточного Мичигана. Из 3453 языков, представленных в Каталоге, 113 распространены в России. Для каждого языка дается его описание; кроме того, размещаются все доступные аудио- и видеоматериалы, указываются действующие проекты по возрождению языка, а также перечисляется основная библиография. Все эти пункты могут быть заполнены любым желающим. Предложенные изменения рассматриваются Консультативной комиссией, состоящей из приглашенных экспертов. К сожалению, информация далеко не по всем языкам представляется нам полной и регулярно обновляемой.

Создатели сайта Глоттолог⁷ также ставили задачу оценить, в какой степени к сегодняшнему дню описаны языки мира (Hammarström et al. 2018). Этот подпроект (или раздел сайта) называется GlottoScore⁸. Для каждого языка указывается “максимальный статус описанности” в зависимости от типа материалов, доступных по этому языку: 1) грамматика больше 300 страниц, 2) грамматика короче 300 страниц, 3) грамматический очерк, 4) текст, описание отдельной языковой особенности, словарь или перевод нового завета, 5) словарный список или меньше. Эту информацию можно визуализировать на карте вместе с информацией о социолингвистическом статусе языков. Такое представление данных дает очень наглядную картину об изученности языков мира в целом. Однако, в отличие от нас, авторы Глоттолога не включали в свой обзор информацию о материалах на языке в дополнение к имеющимся грамматике и словарю. В условиях угрозы исчезновения многих языков России нас в наибольшей степени интересуют материалы на языках, в первую очередь аудио- и видеозаписи.

В ходе работы над составлением перечня языков, требующих срочной документации, и анализа состояния задокументированности языков России нами была

⁶ Сайт проекта: <https://www.endangeredlanguages.com/>.

⁷ Сайт Глоттолог, каталог и классификация языков мира с обширной библиографией: <https://glottolog.org/>

⁸ см. раздел Glottoscope : <https://glottolog.org/langdoc/status>

составлена база данных по миноритарным языкам в виде таблицы “Обзор документации по исчезающим языкам России” (см. Приложение 1). Создание такой базы данных целесообразно ввиду следующих соображений. Во-первых, подобно ресурсу OLAC, описанному выше, информация по имеющимся языковым материалам будет собрана в одном месте. Некоторым преимуществом нашей базы данных является использование русского языка, что сделает ее более доступной для языковых сообществ. Во-вторых, мы объединяем данные по документации и социолингвистические сведения, стараясь указывать возраст младшего носителя и его языковую компетенцию, если такая информация имеется. В-третьих, в базе данных будут указаны недостаточно задокументированные идиомы и рекомендации для работы с конкретными языками. Подобная информация поможет лингвистам не дублировать работу друг друга и расставлять приоритеты для дальнейших исследований, а представителям сообщества предпринимать собственные шаги по документации (см., например, наши рекомендации в Приложении 2).

В идеале, такую базу планируется создать по всем исчезающим языкам России и диалектам, находящимся под угрозой исчезновения, чтобы каждый мог видеть общую картину и знать, какие действия по документации производятся в данный момент. Такая база данных может иметь открытый доступ и может пополняться не только нами, но и другими людьми, имеющими отношение к какому-либо языку (лингвистами, языковыми активистами, представителями языкового сообщества). Для последнего необходимым условием является наличие удобного для пользователей интерфейса с русскоязычной и англоязычной версиями.

Работа по документации языков России имеет давнюю традицию и продолжается в настоящее время многими исследователями. Сбор данных по документации языков показал, что такая общая база актуальна и необходима для нашей страны.

3. Ход работы

3.1. Подготовительный этап

На начальном этапе мы выбрали языки из списка проекта «Языки России» (рук. Ю. Б. Коряков)⁹ с социолингвистическим статусом «Языки на грани исчезновения» и дополнили языками с количеством носителей меньше 1000 (среди них оказались

⁹ Вариант списка 2020 года доступен по ссылке: http://jazykirf.iling-ran.ru/list_concept2020.shtml.

некоторые языки со статусами «Исчезающие языки» и «Языки под угрозой исчезновения с малым числом говорящих»). В ходе работы список был несколько модифицирован с учетом работы группы по уточнению списочного состава языков России. Кроме того, мы дополнили его некоторыми уязвимыми диалектными вариантами в целом более или менее благополучных языков, – это быстринский диалект эвенского, а также тоджинский и терехольский диалекты тувинского языка.

3.2. Разработка параметров

Исходный список языков был составлен по уже известным данным о социолингвистической ситуации языков РФ. Далее необходимо было выбрать, какие еще данные о языках релевантны для решения нашей задачи. Выше (см. 2.3) дано описание тех параметров, которые мы учитывали – остановимся здесь на некоторых нюансах рабочего процесса. Для понимания того, насколько остро язык нуждается в документации, параметр количества носителей необходимо было расширить данными о максимальном уровне владения и возрасте самого младшего носителя. Причем важно учитывать как общее число носителей языка, так и количество носителей для каждого диалекта (особенно в случае, если диалекты достаточно дивергентные). В ходе работы над обзором стало понятно, что нам также нужно по возможности указывать общее количество говорящих и отдельно количество носителей, с которыми можно заниматься документацией.

Параметр объема документации также уточнялся. К грамматикам, словарям и текстам мы добавили разговорники как важную часть документации с точки зрения поддержки и ревитализации языка. Кроме того, мы стали указывать и оценивать отдельно наличие аудио- и видеозаписей хорошего качества. Мы исходим из того, что наличие только письменных материалов в наше время никак не может считаться достаточным. Дополнительным критерием стала доступность имеющихся материалов. Чтобы решить, нуждается ли язык в срочной полевой документации, важно знать, есть ли записи звучащей речи в принципе, независимо от того, полностью ли обработан материал и выложен в сеть Интернет в открытый доступ, или хранится в личных архивах исследователей. С другой стороны, понимать, насколько материал обработан и доступен, важно для планирования актуальных задач по документации в широком смысле. С этой точки зрения также важно, есть ли опыт удаленной работы с носителями

языка. Как мы увидели на воркшопе ELDP по удаленной документации¹⁰, у этого типа работы гораздо больше возможностей, чем мы предполагали. И с каждым годом такое взаимодействие с представителями языковых сообществ становится все более реалистичным.

Кроме того, в ходе работы мы старались найти более подробную информацию о социолингвистической ситуации того или иного языка. Исследования по социолингвистике (как опубликованные статьи, так и информацию о полевой работе по социолингвистике) мы также указывали в индивидуальных таблицах по языкам (см. Приложение 1). В этой связи представляется исключительно важной работа, проводимая в нашем Центре группой, занимающейся оценкой состояния языков России.

3.3. Сбор информации

Базовую информацию, от которой можно было отталкиваться, мы брали из справочных интернет-ресурсов по языкам (основные перечислены выше в 2.5). Некоторые моменты мы уточняли с помощью разнообразных ресурсов, посвященных конкретным языкам. Кроме того, много информации содержится в современной литературе по социолингвистике и документации интересующих нас языков. Про некоторые языки мы могли предоставить детальные данные сами, основываясь на нашем личном исследовательском опыте. Но все же основным источником информации стало личное общение с коллегами, и мы очень благодарны всем, кто смог уделить нам время и поделиться своими знаниями и опытом. Из бесед и переписок мы узнавали наиболее актуальные и детализированные данные. Важным для нас было и субъективное мнение специалистов по каждому параметру. Мы старались его учитывать, заполняя данные по языкам и составляя замечания об актуальных направлениях работы для каждого языка.

На этапе сбора информации стало понятно, насколько это объемная работа. Чтобы составить общее представление о задокументированности всех выбранных языков, нам пришлось отказаться от некоторых деталей. Например, на данном этапе мы не стали описывать каждый диалект в отдельности для всех языков; пришлось отказаться от подробностей о жанровом составе записанных текстов; не получилось указать для всех языков общее число носителей и число свободно владеющих языком (и каждым из диалектов). Не удалось учесть в полной мере и такой фактор, как наличие

¹⁰ <https://elararchive.org/blog/2021/09/03/eldp-remote-fieldwork-workshop-series/>

записанных, но не расшифрованных текстов на языках, что может существенно влиять на планирование работы, ср. [Вахтин, Головки 2005: 49]. Мы планируем постепенно дополнять обзор этой информацией.

При этом и с текущим уровнем детализации хорошо видно, что документация в той или иной степени затронула практически все языки России на грани исчезновения. В то же время стало очевидным, что по каждому языку есть и актуальные задачи в рамках документации. Далее в разделах 4 и 5 представлены краткие выводы о документации каждого языка. Более подробно с собранной нами информацией можно ознакомиться в таблице “Обзор документации по исчезающим языкам России”.

4. Языки или диалекты, для которых наиболее актуальна документация в 2022 году

Языки разбиты на группы в соответствии с той задачей, которая для них приоритетна. Это, конечно, не означает, что других задач не стоит. Для языков первой группы актуальна собственно полевая документация; при этом можно считать, что социолингвистическая ситуация с ними вполне ясна. Для языков второй группы актуален сбор социолингвистических данных, а по полученным данным можно будет планировать дальнейшие действия в рамках документации.

4.1. Языки, для которых актуальна собственно полевая документация

Ительменский

Из языков этой группы ительменский язык, пожалуй, задокументирован больше всего. Для ительменского языка есть понимание социолингвистической ситуации, есть грамматические описания, опубликованные тексты, озвученный словарь, видеозаписи текстов. В то же время необходимо дополнять имеющиеся материалы (не только количественно, но и качественно, например, записывать повседневные фразы и диалоги, речь, обращенную к детям). При этом возможность полевой документации в ближайшие годы может исчезнуть. Сейчас ительменским языком на уровне, достаточном для полноценной документации, владеет только несколько очень пожилых человек 1930-х годов рождения, проживающих в пос. Тигиль и Седанка. В других поселках также проживает меньше 10 человек, в разной степени владеющих языком. Важно, что в самом языковом сообществе есть интерес к судьбе языка, есть активисты, занимающиеся

поддержкой и сохранением языка. Кроме того, активной полевой документацией ительменского языка российскими исследователями сейчас не ведется. Однако, на базе ВШЭ и “Лаборатории исследования и сохранения малых языков России” (ИЯз РАН, Москва) сформирована группа исследователей (рук. К. О. Шейфер), которая с 2019 года работает с уже собранными материалами, создала корпус ительменских текстов и планирует экспедицию на Камчатку.

Амурский нивхский: северносахалинский диалект

По нивхским языкам собраны довольно представительные языковые материалы, однако по северносахалинскому диалекту нет ничего: ни грамматики, ни словаря, ни текстов. В то же время владеющих этим диалектом сейчас меньше 10 человек. При этом в целом нивхское языковое сообщество активно вовлечено в поддержку своего языка, и в настоящее время ведется проект по ревитализации языка. В такой ситуации особенно важно иметь аудио- и видеоматериалы на каждом диалекте, так как диалектные различия необходимо учитывать при создании учебных материалов. Такие материалы возможно собрать в ходе документации. Кроме того, поскольку нивхские языки считаются генетически изолированными, северносахалинский диалект представляет исключительный интерес с точки зрения исторического развития этих языков.

Южноюкагирский

Южноюкагирский язык хорошо описан, есть академическая грамматика, посвященная именно ему (см. [Maslova 2003]). Издано много учебной литературы и словарей. Непосредственно документацией южноюкагирского занималась Ирина Алексеевна Николаева. На ее сайте¹¹ можно найти 52 текста разных жанров с аудиозаписями и глоссами. Таким образом, для южноюкагирского необходимо записать полноценный озвученный словарь, дополнить коллекцию текстов видеозаписями и такими жанрами, как повседневные фразы и диалоги, и речью, обращенной к детям. В настоящее время активной документацией южноюкагирского не ведется, при этом владеющих языком на достаточном для этого уровне осталось единицы. К тому же в сообществе есть люди, активно интересующиеся судьбой языка.

¹¹ Сайт доступен по следующей ссылке: <https://www.sgr.fi/yukaghir/index.html>.

Нешанский: аганский говор

Нешанский (или лесной ненецкий) язык хорошо описан, изданы словари, опубликованы тексты. В нешанском выделяются три говора: нумтовский, пуровский и аганский, которые отличаются незначительно фонетически и лексически. Все три говора в некоторой степени задокументированы. В то же время владеющие аганским говором, с кем можно было бы вести документацию, – это максимум несколько десятков человек пожилого возраста. Наиболее актуальной задачей является запись озвученного словаря. Кроме того, необходимо пополнить коллекцию видеозаписями и текстами таких жанров, как бытовые диалоги и повседневные фразы.

4.2. Языки, для которых актуально социолингвистическое исследование

Науканский

Науканский язык, ранее считавшийся диалектом единого языка азиатских эскимосов, отдельно описан в труде Г. А. Меновщикова, там же приводятся тексты с переводом [Меновщиков 1975]. Кроме этого, издано два академических словаря [Fortescue et al. 1994; Головкин и др. 2004]. В остальном материалов по языку фактически нет (за исключением отдельных статей). Одновременно нет и актуальной информации о числе владеющих языком на том или ином уровне. Наиболее свежие данные о приблизительном количестве носителей языка приводятся в [Jakobson 2005] — исследователь дает оценку в менее чем 50 человек. Таким образом, необходимо в ближайшее время провести социолингвистическое исследование и документацию в той степени, в какой это позволят сделать языковые компетенции носителей.

Хамниганский

По грамматике ононского хамниганского (варианта хамниганского языка, на котором говорят в России) изданы очерки и отдельные статьи. В 2015 году вышел хамниганско-русский словарь (см. [Сундуева 2015]). Несмотря на то, что в регионе проводятся социолингвистические исследования (см., к примеру, [Сундуева 2010]), ситуация до конца не ясна. Можно быть уверенными, что свободно владеющие языком хамниганцы есть, но их количество неизвестно. При этом аудио- и видеоматериалов именно ононского варианта на данный момент нет, что означает необходимость полноценной документации данного идиома. Общественный интерес к судьбе языка в сообществе есть. Так, например, известно, что по поручению Главы Буддийской

традиционной сангхи России в селе Курульга уже много лет составляется хамниганский словарь.

Сойотский

Большинство материалов по сойотскому языку – работа Валентина Ивановича Рассадина: грамматические очерки и статьи, словари, учебные пособия. Валентин Иванович создал письменность для сойотского и планировал проект по ревитализации языка. Однако аудио- и видеозаписей речи нет совсем. Сойотский язык считается заснувшим, при этом ведется его преподавание. Сейчас срочной представляется задача выяснить, есть ли владеющие сойотским языком, каков уровень владения, и по возможности записать языковой материал (звук и видео).

Сирхя-цудахарский: амухский и анклухский

Языки Дагестана находятся не в столь критическом положении, как описанные выше, однако в то же время процесс языкового сдвига в регионе местами идет очень активно, а по многим идиомам полноценной документации не проводилось и не проводится. Кроме того, некоторые идиомы мало изучены в принципе. Среди таких случаев в ходе консультаций со специалистами как претенденты на работу по документации нами были выделены близкие амухский и анклухский варианты сирхя-цудахарского языка. По ним есть отдельные статьи И. А. Курбанова, собранный им же 200-словник и несколько коротких аудиозаписей, но в остальном материалов (в том числе аудио и видео) нет. О числе владеющих языком точных данных также нет, но предположительно оно исчисляется десятками. Так как амухцы и анклухцы переезжают из родных горных селений (Амуха и уже опустевшего Анклуха) в долину, в многонациональные поселки и города, языковая ситуация очень нестабильна, что является дополнительным фактором для проведения документационной и социолингвистической работы с этими идиомами.

Южнорутульский

Рутульскими языками активно занимаются разные группы лингвистов. В то же время именно южнорутульский (или борчинско-хновский) описан недостаточно. Нужно, однако, отметить достаточно полный словарь, вышедший в 2021 году [Khalilov 2021]. Аудио- и видеоматериалов, отвечающих современным стандартам документации, практически нет. При этом язык, судя по всему, находится под угрозой исчезновения.

По данным экспедиций ВШЭ в этот регион, из села Хнов жители активно уезжают. Что касается села Борч, то его можно считать вымершим: его жители предположительно частично переселились в село Новый Борч, про языковую ситуацию в котором известно крайне мало. Таким образом, помимо документации необходимо также социолингвистическое исследование.

4.3. Языки, документация которых проводится в 2021 г. на базе Научного центра в качестве пилотного проекта

Уильтинский

Носителей уильтинского (или орокского) языка осталось менее 10 человек пожилого возраста. Из них только два-три человека владеют языком свободно. Материалов по уильтинскому собрано достаточно много, есть грамматические описания и словари на обоих диалектах (северном и южном), вопросами грамматики сейчас занимается аспирант П. Червинский (Университет Майнца, Германия). Однако ситуация с документацией уильтинского показалась нам достаточно срочной. Хотя работа по сбору и расшифровке материала велась разными группами исследователей, на сегодняшний день некоторые материалы были не расшифрованы, а многие другие — не отгlossированы. Для публикации коллекции текстов в Интернете требуется дальнейшая обработка материалов, в частности, синхронизация транскрипции со звуковыми и видеофайлами. Мы получили материалы, имеющиеся у других ученых (В. Ю. Гусева, О. А. Казакевич), а также продолжили работу в проекте FieldWorks, в котором уже отгlossировано около 60 текстов. Носители уильтинского активно публикуют различные материалы на своем языке (в основном этим занимается Е. А. Бибилова). С большим энтузиазмом была воспринята идея озвучить уильтинский разговорник, составленный Е. А. Бибиловой. Во время полевой работы в г. Поронайске мы восполнили некоторые жанровые пробелы в собранных ранее материалах — нами были записаны процедурные тексты (обработка рыбьей кожи, традиционные игры) и бытовая речь. Кроме того, мы работали над расшифровкой старых и новых материалов, а также записывали звуковой материал для планируемого разговорника.

5. Обзор других идиомов

5.1. Языки, для которых актуально социолингвистическое исследование

Нами было выделено семь идиомов, для которых было бы актуально проведение социолингвистического исследования. Выбранные идиомы можно условно разделить на две группы. Для первой группы, куда мы включаем крымчакский и калмакский языки, а также тазовский диалект китайского и крымский диалект караимского, социолингвистическое исследование является единственной актуальной задачей, поскольку, насколько известно, сейчас для идиомов этой группы нет носителей, с кем можно было бы заниматься полноценной полевой документацией, и одновременно с этим по ним нет почти никаких материалов, с которыми можно было бы работать. В то же время мы предлагаем рассмотреть для них возможность “фрагментарной” документации, то есть фиксацию звучащей речи, насколько это позволит сделать компетенция тех, кто все же в какой-то степени владеет языком (что и выяснится при проведении социолингвистического исследования). Для второй группы идиомов — алабугатско-ногайского и тубаларского языков, а также томского диалекта сибирскотатарского языка — также актуально социолингвистическое исследование, однако, в отличие от первой группы идиомов, с ними также возможно и проведение работы по документации, так как, как мы полагаем, имеются полноценные носители данных идиомов.

Крымчакский

Материалов по крымчакскому языку немного, но все же есть грамматические описания, словарь и опубликованные тексты. В то же время полноценной документации не проводилось. В 1980-х годах материал собирала Мария Семеновна Полинская (8 текстов с поморфемным разбором опубликованы в [Polinsky 1991]). Уже в то время крымчаки младше 70 лет не знали этнического языка. Больше документацией крымчакского никто не занимался.

Караимский: крымский диалект

В целом по крымскому диалекту караимского языка материалы имеются, но их недостаточно, нужны звуковые материалы, нет современных грамматических описаний и мало текстов (в наличии, насколько нам известно, имеется только сборник

музыкального фольклора крымских караимов [Кефели 2018]). Информация по количеству носителей противоречива, поэтому необходимо тщательное социолингвистическое исследование данного идиома. В. А. Ельяшевич считает крымский караимский мертвым этнолектом крымскотатарского языка. Однако, нам стало известно, что один из последних носителей крымского караимского использует язык в литургии. По нашим сведениям, материалы экспедиции РГГУ 1998 г. могут храниться у М. С. Полинской. Возможно, анализ этих материалов – потенциальное направление для документации этого языка.

Китайский: газовский диалект

В Институте языкознания хранится архив И. В. Самариной с материалами на газовском диалекте китайского языка. Инвентаризация этого архива и работа с его материалами планируется в рамках Лаборатории исследования и сохранения малых языков. Других материалов нет. В этом году Марией Яковлевной Каплуновой были записаны отрывочные образцы речи. По ее данным свободно на этом языке уже никто не говорит. При этом есть носители, которые могут произнести отдельные фразы и слова, но в то же время они совершенно не заинтересованы в работе с лингвистами, закрываются при просьбе произнести что-то “на своем” и вообще не любят эту тему. В такой ситуации можно провести социолингвистическое исследование, чтобы более четко обрисовать языковую ситуацию. По возможности надо записать то, что получится, чтобы сохранить хоть какие-то примеры звучащей речи. При проведении систематической социолингвистической работы с данным идиомом отношение сообщества к языку может измениться.

Калмакский

По калмакскому языку есть отдельные грамматические статьи, социолингвистические исследования последних лет, а также работы по фонетике Николая Сергеевича Уртегешева. В открытом доступе имеются озвученные словарные материалы на платформе Лингводок. По личному сообщению Н. С. Уртегешева, свободно владеющих языком сейчас нет. В то же время можно подумать о том, чтобы сделать коллекцию примеров звучащей речи, которые еще, может быть, удастся записать с теми, кто помнит отдельные фразы и слова на калмакском.

Алабугатско-ногайский

Для алабугатско-ногайского очень актуальна полевая документация, так как аудио- и видеоматериалов фактически никаких нет. Информации об этом языке крайне мало. Но в первую очередь необходимо масштабное социолингвистическое исследование, которое показало бы, где проживают носители именно алабугатско-ногайского (так как существует проблема разграничения данного идиома с ногайским и юртовско-ногайским), а также их численность и языковые компетенции. По предварительным оценкам, число владеющих алабугатско-ногайским в разной степени составляет от 400 до 1000 человек.

Сибирскотатарский: томский диалект

По томскому диалекту сибирскотатарского языка материалов не так много, однако исследования ведутся. Есть районы предположительного проживания носителей этого диалекта, где не проводилось социолингвистического исследования. Число владеющих томским диалектом сибирскотатарского необходимо уточнять, как и языковую ситуацию в целом. Кроме того, по возможности следует проводить документацию.

Тубаларский

Тубаларский язык хорошо описан, но нуждается в документации – материалов пока собрано мало, в основном они находятся у разных исследователей в личных архивах. Несмотря на то, что языковая ситуация пока не столь критическая, наиболее актуальным на сегодняшний день представляется детальное социолингвистическое исследование. В то же время можно расширять коллекцию видеозаписей текстов и записывать более полные озвученные словари.

5.2. Языки, для которых актуальна работа с имеющимися данными

Языки этой группы были задокументированы в разной степени, однако не весь массив собранных ранее материалов был обработан и стал доступен общественности. При этом в настоящее время продолжение полноценной документации этих языков невозможно либо крайне маловероятно.

Бабинский саамский

Для бабинского саамского есть грамматическое описание, словарь и записанные тексты. Тексты хранятся сейчас в архивах. Таким образом, актуальны работа с архивными материалами и представление их в сети Интернет. По словам исследователей, носителей бабинского саамского, с которыми можно было бы заниматься полноценной документацией, сейчас уже нет. Однако в России живут две женщины, помнящие язык, с которыми можно попробовать записать то, что они помнят (звук и видео). Кроме того, в сообществе бабинских саамов проявляется интерес к судьбе языка.

Йоканьгско-саамский

Для йоканьгско-саамского опубликовано довольно много материалов: есть грамматика, статьи на отдельные темы, словарь, тексты с переводом. Некоторые аудио- и видеоматериалы тоже есть (они хранятся в архивах, а один словарь представлен на платформе Лингводок). И хотя материалов явно недостаточно, реализуемость проекта по документации находится под вопросом, так как, по данным исследователей, осталась одна носительница йоканьгско-саамского языка. Поэтому нам представляется важным провести работу с архивными материалами и представить их в открытом доступе.

Медновский алеутский

С медновским алеутским социолингвистическая ситуация представляется достаточно ясной. Что касается документации, то по медновскому алеутскому имеются грамматические описания, издан словарь. Сейчас языком владеют два пожилых человека, которые уже не могут работать с лингвистами. Достаточно много материала по медновскому алеутскому (как и по беринговскому диалекту алеутского языка) собрал Евгений Васильевич Головкин; в настоящее время он продолжает работу с этими идиомами.

Водский

По водскому языку издано довольно много материалов: есть грамматики, грамматические статьи, тексты (в том числе с поморфемной аннотацией), в меньшей степени словари. Велись многолетние документационные проекты нескольких групп исследователей: Ф. И. Рожанского и Е. Б. Маркус, М. З. Муслимова, Энна Эрнитса, Н.

В. Кузнецовой, группой из Карельского научного центра РАН. Коллекция текстов, собранных перечисленными группами исследователей, представлена в архиве ELDP¹². Объем аудиоматериалов составляет 83,7 часа, видеоматериалов — 6,2 часов. По словам Ф. И. Рожанского, носителей, свободно владеющих языком, на сегодняшний день не осталось. В то же время довольно активно ведутся проекты по ревитализации языка. В таком контексте особо актуально представление собранных данных в таком виде, чтобы их можно было использовать для целей сохранения языка, например, для создания учебных материалов.

5.3. Языки, для которых актуальна работа с имеющимися данными и пополнением материалов

Нами была выделена группа достаточно хорошо задокументированных языков, для которых актуальны и обработка и обеспечение доступности собранных материалов, и продолжение полевой документации. Тем не менее, приоритетной среди этих двух задач мы считаем работу с уже собранными языковыми данными.

Негидальский

Для негидальского языка есть грамматические описания и статьи, издан словарь, опубликованы тексты с переводом. Кроме того, за последние 15 лет были проведены два больших документационных проекта (под рук. Е. Ю. Калининой и под рук. Б. Пакендорф), в ходе которых было сделано много аудио- и видеозаписей текстов разных жанров и был записан звучащий словарь. Собранные материалы можно считать достаточными, хотя их можно было бы дополнить записью бытовой коммуникации. На данный момент актуальна задача обработки собранных материалов и представления их в сети Интернет (в открытом доступе как для лингвистов, так и для сообщества). В настоящее время лишь несколько человек владеют негидальским языком, и все они в последние годы уже работали с лингвистами.

Тофаларский

В целом язык хорошо описан (по большей части В. И. Рассадиным). С 2007 года ведется наблюдение над состоянием языка А. А. Сюрюн. В 2018 году была проведена полноценная документация языка; аудио- и видеоматериалы сейчас хранятся в архиве

¹² Ссылки на соответствующие депозитарии см. в Приложении 1 для конкретных языков.

ELDP. Материал для словаря и разговорника собран. А. А. Сюрюн продолжает активную работу с тофаларским языком, в том числе направленную на поддержку языка, в том числе обеспечение открытого доступа ко всем языковым материалам, что сейчас особенно актуально, и поддерживает контакт с носителями языка.

Удэгейский

Удэгейский язык находится в критической ситуации, осталось около 10 носителей. Материалов по нему достаточно много. Есть и грамматические описания, и тексты (в том числе с поморфемным разбором), и словари. Кроме того, собрано представительное количество аудио- и видеоматериалов, которые частично доступны в сети Интернет. На данный момент актуальной представляется задача обеспечения доступности собранных материалов. Последние года документацией и поддержкой удэгейского языка активно занимались Е. В. Перехвальская, И. А. Николаева и М. В. Тольская.

Центрально-южноселькупский

По центрально-южноселькупскому языку собрано достаточно много документационных материалов, имеются словари (в том числе и озвученный), тексты, аудио- и видеозаписи. Тем не менее, полноценное описание грамматики данного языка отсутствует, имеются лишь отдельные статьи. Носителей, с кем можно было бы продолжать документацию, остались единицы. При этом в сообществе есть попытки поддержки и ревитализации языка. Таким образом, актуальна дальнейшая обработка собранных данных и представление их в открытом доступе, чтобы их можно было использовать для сохранения языка.

Скольтский саамский

По скольтскому саамскому языку издано довольно много грамматической литературы, но мало словарных материалов, нет публикаций текстов. Аудиоматериалы (словари и тексты) можно найти на платформах с разным уровнем доступности (Лингводок, DOBES), а также в архиве Карельского научного центра. По словам М. Рисслера, вряд ли в документации нотозерского диалекта скольтского саамского (на котором говорят на территории России) можно сделать что-то кардинально новое. Кроме того, в Финляндии есть носители другого варианта скольтского, который близок распространенному в России, и с ними также работают лингвисты. Актуальное

направление на данный момент – это работа с архивными материалами, представление их в открытом доступе. Это важно еще и потому, что языковое сообщество заинтересовано в этих материалах, чтобы их можно было использовать для сохранения языка.

Лесной энецкий

Лесной энецкий язык довольно полно описан, есть много опубликованных материалов. Владющих языком, кто мог бы работать с лингвистами, сейчас только несколько десятков. Документация и мониторинг социолингвистической ситуации активно ведутся А. Б. Шлуинским, О. В. Ханиной и М. А. Овсянниковой. Собрано достаточно аудио- и видеоматериалов. При этом они доступны лишь частично, что делает задачу представления их в открытом доступе актуальной.

Тундровый энецкий

Тундровый энецкий язык достаточно хорошо описан и задокументирован, при этом собранные материалы лишь частично доступны. Носителей тундрового энецкого языка осталось крайне мало, все они пожилые люди. Тундровым энецким активно занимаются разные группы ученых, однако важно работать в направлении доступности материалов для представителей сообщества.

Нганасанский

Нганасанский язык достаточно хорошо описан, по нему издано много материалов, в том числе учебных. Владющих языком, с кем можно было бы заниматься полноценной документацией, осталось несколько десятков, но при этом уже собрано представительное количество аудио- и видеоматериалов (тексты и словари). Большая часть материалов доступна для исследователей. В то же время у сообщества есть активный запрос на поддержку и ревитализацию языка. Таким образом, наиболее актуальной задачей в рамках документации сейчас является представление собранных материалов в доступном для сообщества виде, чтобы их можно было использовать для создания учебных материалов.

Чулымский

По чулымскому языку издано достаточно много описаний, ведется работа над объемным словарем, есть опубликованные тексты (в том числе с поморфемным

разбором). Сейчас ведется активная документация среднечулымского диалекта, в ходе которой все носители, владеющие языком на достаточном уровне, работают с лингвистами. Уже собрано достаточно много материалов, которые частично доступны, а частично находятся в личных архивах, в архивах ELDP и ТГПУ. Важно отметить, что также задокументирован и нижнечулымский диалект, который с 2011 года уже является заснувшим. Таким образом, на данный момент обработка собранных чулымских материалов представляется особенно важной.

Алюторский

Язык в целом достаточно хорошо описан, при этом карагинский диалект представлен в разрозненных изданиях, а об укинском диалекте известно крайне мало. Есть изданные тексты (большой частью с переводом, но есть с поморфемным разбором). Есть коллекция текстов со звуком и видео в архиве ELDP. Однако нет онлайн коллекции текстов, доступной для сообщества. При этом передача языка детям прервана уже довольно давно. На данный момент свободно владеют языком люди старше 50 лет, составляющие около 10% этнической группы. Людей младше 40 лет, говорящих на языке, – единицы. По оценкам специалистов, общее число носителей не превышает 150 человек. Таким образом, для алюторского языка актуальны продолжение документационных проектов (с координацией усилий разных исследователей) и обработка собранных материалов (создание корпуса, создание онлайн ресурсов для сообщества). Стоит отметить, что в последние годы российских исследователей алюторского нет.

Финский: ингерманландский диалект

По ингерманландскому диалекту финского языка изданы грамматические статьи, есть диссертации по отдельным грамматическим темам, один словарь, одна публикация текстов. В Эстонии сейчас создается большой онлайн-словарь. В последние годы собрано много текстов разных жанров на всех говорах (в том числе М. З. Муслимовым, который активно занимается исследованием ингерманландского финского). Материалы сейчас находятся в личном архиве М. З. Муслимова и в архиве Kotus. Оценить количество владеющих ингерманландским диалектом, проживающих на территории России, на данный момент очень сложно: во-первых, многие из них люди преклонного возраста, во-вторых, ингерманландцы активно уезжают в Финляндию или крупные города России. Учитывая объем и разнообразие собранных материалов, представляется

актуальной их обработка и создание доступных ресурсов (как для лингвистов, так и для сообщества).

Кетский

Кетский язык достаточно хорошо описан и задокументирован. Изданы грамматики и множество грамматических статей, словари (в том числе учебные), тексты (в основном с переводом). Есть немного школьной литературы. Записано достаточно много аудио- и видеоматериалов (словарей и текстов) от носителей разных диалектов. На данном этапе актуально записывать тексты недостающих жанров (бытовые диалоги, повседневные фразы, речь, обращенная к детям), а также обрабатывать и делать доступными сообществу собранные ранее материалы (в том числе архивные).

Северноюкагирский

По сравнению с южноюкагирским, изучен и задокументирован лучше. Для более полной документации не хватает озвученного словаря. При этом северноюкагирским активно занимаются разные группы ученых в России и за рубежом. По нашим сведениям, этот язык будет включен в планируемый проект по документации языков Нижней Колымы на базе университета Мюнстера (Германия).

Ижорский

По ижорскому языку имеется обширная документация: записаны как видео-, так и аудиоматериалы, есть озвученный словарь. На данный момент значительная часть материалов доступна для лингвистов. Активная работа ведется группой Ф. И. Рожанского и Е. Б. Маркус, М. З. Муслимовым, Т. Б. Агранат и другими исследователями. Языковое сообщество заинтересовано в судьбе языка, поэтому представление собранных материалов в доступной форме очень актуально.

Эвенский: быстринский диалект

По быстринскому диалекту эвенского языка было собрано достаточно много документационных материалов в проекте "Dialectal and cultural diversity among Evens in Siberia" при Институте эволюционной антропологии Общества Макса Планка. В настоящее время в Быстринский район Камчатского края регулярно осуществляются экспедиции студентов и сотрудников НИУ ВШЭ. Для быстринского диалекта целесообразна видеозапись текстов недостающих жанров, а также запись звучащего

словаря. Кроме того, необходима обработка материалов с учетом запроса и активной деятельности по поддержке языка.

Кильдинский саамский

Кильдинским саамским языком интенсивно занимался проф. М. Рисслер. Им записано довольно много аудиоматериалов, которые хранятся в ТЛА. По кильдинскому саамскому языку есть в том числе корпус и коллекция отгlossированных текстов. Численность носителей кильдинского саамского оценивается в несколько сотен. В сообществе есть группа людей, равнодушных к судьбе языка. Сейчас наиболее актуальной в рамках документации кажется задача обработки существующих материалов с прицелом на ревитализацию языка.

Шорский

Документацию шорского языка активно ведут исследователи из различных городов и стран (в частности, из Новокузнецка, Новосибирска, Франкфурта, Москвы). По шорскому языку есть корпус текстов и небольшой озвученный словарь. По информации, полученной от Д. М. Токмашева, количество носителей составляет около тысячи человек, однако уверенных носителей всего лишь несколько сотен. Межпоколенческая передача нарушена около двадцати лет назад. В такой ситуации актуально использовать собранные материалы для поддержки и сохранения языка, а также публиковать материалы из личных архивов.

Телеутский

По телеутскому языку издано довольно много литературы (описания, словари, тексты, учебная литература). Документация на данный момент не очень обширная, но ведется работа Д. М. Токмашевым и новосибирскими лингвистами. Количество носителей телеутского языка в настоящее время оценивается сотнями, однако межпоколенческая передача языка нарушена. Важно отметить, что ведется поддержка языка, предпринимаются попытки его ревитализации. Поэтому обработка и представление материалов в открытом доступе являются крайне важными задачами для этих целей.

Корякский

По последним оценкам, носителей корякского языка осталось несколько сотен. По языку изданы академическая грамматика, словари, опубликованы тексты. Есть аудиоматериалы в архиве ELDP, собранные Эрихом Кастеном и Александром Кингом. В последние годы в места компактного проживания коряков проводятся экспедиции студентов и сотрудников НИУ ВШЭ. Для корякского языка актуальна задача представления собранных данных в открытом доступе, а также пополнение материалов.

5.4. Языки, для которых актуальна собственно документация

Языки этой группы нуждаются в документации, однако они не включены в список в п. 4.1. Такое решение связано либо с тем, что носителей этих языков значительно больше, чем носителей языков, перечисленных в п. 4.1, либо с тем, что документация языка уже активно ведется другими исследователями.

Ульчский

По грамматике ульчского языка есть очерки и статьи на отдельные темы, изданы учебные словари и пособия, опубликованы тексты с переводом, есть также записанные тексты и в личных архивах. Кроме того, имеются озвученные словари по разным диалектам. Тем не менее, объем документационных материалов представляется явно недостаточным. При этом носителей ульчского осталось только около 40 человек пожилого возраста. В настоящее время есть активные проекты по документации этого языка.

Чаплинский

По чаплинскому языку опубликовано достаточно много научной литературы, имеется грамматика этого языка и словарь. Кроме того, на чаплинском языке издана и художественная литература. Тем не менее, полноценной документации чаплинского языка на территории России не проводилось, однако некоторые аудиоматериалы все же доступны. Для чаплинского актуально расширять документационные проекты, а также держать фокус на поддержку языка, поскольку мнение сообщества в последние года меняется, появляется интерес к судьбе языка. Владеющих языком на том или ином уровне сейчас осталось около 100 человек старше 40 лет. Стоит отметить, что вариант

чаплинского языка, распространенный на о. Св. Лаврентия, довольно хорошо задокументирован американскими лингвистами.

Сето

Информации по языку сето имеется не так много. Однако работа в этом направлении ведется (как с вариантом Псковской области, так и с языком сибирских сето), есть проекты по поддержке и ревитализации языка. Владеющих языком сейчас больше ста человек. Для языка сето наиболее актуальным представляются продолжение документации и пополнение аудио- и видеоматериалов.

Кумандинский

Свободно владеющих кумандинским языком, по оценкам специалистов, в настоящий момент несколько сотен. По кумандинскому языку есть несколько грамматических описаний, словарей и разговорников. В архиве ИФ СО РАН хранятся аудиозаписи текстов на кумандинском. Однако в остальном материалов по этому языку нет, поэтому на данный момент степень его задокументированности можно считать недостаточной. В ближайшие 5 лет работа с кумандинским языком запланирована А. В. Дыбо и ее командой.

Кантыкский: васюганский диалект

По кантыкскому (или восточнохантыйскому) языку в целом ведется достаточно много работы, по нему имеются словари и тексты. В восточнохантыйском языке выделяются сургутский и васюганский диалекты, при этом носителей последнего остались единицы. Документацию васюганского диалекта необходимо расширять, и в настоящее время имеются исследователи, которые работают в этом направлении. С лингвистической точки зрения отличия между двумя диалектами кантыкского небольшие, однако носителями они воспринимаются как существенные. Это важно с точки зрения создания учебных материалов, базой для которых могут явиться собранные документационные материалы..

Нанайский: горинский диалект

Горинский диалект нанайского языка довольно сильно отличается от остальных диалектов нанайского, что делает документацию особенно актуальной. Говорящих на этом диалекте, с которыми можно было бы работать, предположительно несколько

десятков человек. К настоящему моменту есть собранный звучащий словарь и некоторое количество записанных текстов. Работа по документации горинского диалекта ведется Н. М. Стойновой и другими исследователями.

Тувинский: тоджинский и терехольский диалекты

По данным диалектам собственно документационная работа не проводилась. Вероятно, это связано с тем, что носителей данных диалектов достаточно много; отмечается, что в некоторых местах еще относительно хорошо сохраняется межпоколенческая передача языка. Аудио- и видеоматериалов по данным диалектам очень мало, и все они находятся в личных архивах исследователей. Задача документации представляется актуальной (как и обработка имеющихся материалов), но не кажется срочной ввиду достаточно большого числа носителей.

Вепсский

По вепсскому языку издано много материалов, в том числе художественная и учебная литература. Количество носителей вепсского языка составляет несколько тысяч, однако при этом межпоколенческая передача нарушена. В плане документации вепсского языка есть пространство для деятельности — это и запись озвученных словарей, и осуществление видеозаписей, и обработка имеющихся материалов.

Кроме того, необходимо уточнить социолингвистическую ситуацию и объем документации по разным диалектам вепсского. В настоящее время исследованием вепсского языка занимаются несколько групп исследователей.

6. Заключение

Наше исследование по оценке документации исчезающих языков России показало, что какой-то объем работы по документации осуществлен по большинству рассмотренных языков. В то же время, во многих случаях эта документация неполная (например, часто отсутствуют видеоматериалы или озвученный словарь) и не отвечает современным требованиям относительно технических характеристик, доступности материала и актуальности описания, социолингвистического исследования языка и его диалектов. Выделенные в разделе 4 языки нуждаются в наиболее активных действиях по документации. Однако и для многих языков, описанных в разделе 5, также нужна дополнительная работа по разным направлениям документации. К сожалению,

дальнейшая работа возможна не для всех перечисленных языков из-за отсутствия носителей языка. Конкретные направления работы мы указали в кратком описании каждого языка. В целом наиболее важными направлениями нам видятся полевая работа с языками, для которых это еще возможно, публикация имеющихся материалов и организация их на специализированных платформах в Интернете, а также дальнейшая работа по анализу собранного материала. Наш обзор еще раз подтвердил, что сама запись материала – это, безусловно, очень важный шаг в работе по лингвистической документации, но именно публикация аннотированного материала, его представление в форме, доступной для академического сообщества, самих носителей и их потомков, делает эту документацию полноценным источником информации о языке.

В процессе сбора информации мы консультировались с большим количеством специалистов по разным языкам. Многие из них выражали заинтересованность в нашей работе и готовность к сотрудничеству в той или иной форме. Специалисты по тюркским языкам, языкам Кавказа, саамским языкам признавали проблему нехватки кадров для документации и необходимость в координации действий исследователей. Создаваемая база данных станет удобным инструментом для такой координации, востребованной и значимой как для академической лингвистики, так и для практических целей поддержки и ревитализации языков.

7. Благодарности экспертам

Мы хотим выразить особенную признательность всем специалистам, которые делились с нами знаниями о социолингвистической обстановке в разных регионах и накопленных материалах по тем или иным языкам. За прошедшие месяцы с нами разговаривали и переписывались, отвечали на многочисленные вопросы следующие эксперты: Л. Ф. Абубакирова, Ф. Г. Алексеев, А. В. Архипов, Т. Б. Агранат, В. Б. Бакула, И.С. Буркова, Н. Б. Вахтин, Е. В. Головкин, И. А. Грунтов, В. Ю. Гусев, Н. Р. Добрушина, А. В. Дыбо, В. А. Ельяшевич, М. М. Зимин, Р. В. Ишмухамбетов, С. Казама, О. А. Казакевич, Е. Л. Клячко, Ю. Б. Коряков, Н. В. Кропотова, Ю. А. Ландер, Т. А. Майсак, Д. Н. Моргунова-Швальбе, А. Д. Мусаев, М. З. Муслимов, Р. О. Муталов, И. Нагасаки, Ю. Нагаяма, Ч. Оно, С. А. Оскольская, Д. Т. Пиркова, М. Ю. Пупынина, Ф. И. Рожанский, Д. С. Сатанова, Л. С. Селендили, В. В. Совкина, Н. М. Стойнова, Н. Р. Сумбатова, Я. Г. Тестелец, Д. М. Токмашев, С. Ю. Толдова, Н. С. Уртегешев, М. С. Федина, А. Ю. Фильченко, К. О. Шейфер, Ю. Янхунен.

Литература

Вахтин Н. Б., Головки Е. В. Исчезающие языки и задачи лингвистов-североведов. // Малые языки и традиции: существование на грани. Вып. 1. Лингвистические проблемы сохранения и документации малых языков. Посвящается 75-летию академика Вячеслава Всеволодовича Иванова / Под редакцией А. Е. Кибрика – М.: Новое издательство, 2005. С. 40–52.

Головки Е. В., Джейкобсон С., Добриева Е. А., Краусс М. Словарь языка науканских эскимосов. Ок. 6 тыс. слов, с указателем суффиксов и списком топонимов. Фэрбенкс: Университет штата Аляска в Фэрбенксе, Центр изучения языков коренного населения, 2004.

Кефели А. Сборник музыкального фольклора крымских караимов. Симферополь: ИП Зуева Т. В., 2018.

Меновщиков Г. А. Язык науканских эскимосов. Л.: Наука, 1975.

Сундуева Е. В. Агинское Приононье как социолингвистический феномен. Вестник ЧитГУ. 5(62). 2010. С. 50–57.

Сундуева Е. В., Дамдинов Д. Г. Хамниганско-русский словарь. Иркутск: Изд-во "Оттиск", 2015.

Amery R. Phoenix or relic? Documentation of languages with revitalization in mind. *Language Documentation & Conservation* 3 (2). 2009. Pp. 138–148.

Austin P. K. Language documentation 20 years on // Filipović L., Pütz M. (eds.). *Endangered Languages and Languages in Danger. Issues of documentation, policy, and language rights*. Amsterdam: John Benjamins, 2016. Pp. 147–170.

Austin P. K., Sallabank J. Language documentation and language revitalization: Some methodological considerations // Hinton L., Huss L., Roche G. (eds.). *The Routledge handbook of language revitalization*. London: Routledge, 2018. Pp. 207–215.

Fortescue M., Jacobson S., Kaplan L. *Comparative Eskimo Dictionary with Aleut Cognates*. Fairbanks, Alaska Native Language Center, Research paper, 9. 1994.

Hale K., Krauss M., Watahomigie L. J., Yamamoto A. Y., Craig C., Jeanne L. M., England N. C. *Endangered languages*. *Language* 68(1). 1992. Pp. 1–42.

Harald Hammarström and Thom Castermans and Robert Forkel and Kevin Verbeek and Michel A. Westenberg and Bettina Speckmann. 2018. Simultaneous Visualization of Language Endangerment and Language Description. *Language Documentation & Conservation* 12. 359–392.

Jacobson S. A. History of the Naukan Yupik Eskimo dictionary with implications for a future Siberian Yupik dictionary. *Études Inuit Studies* 29(1-2). Pp. 149–161.

Khalilov M. Rutul (Borchino Khnow dialect) dictionary // Key M. R., Comrie B. (eds.). *The Intercontinental Dictionary Series*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2021.

Maslova E. *A Grammar of Kolyma Yukaghir*. Berlin, New York: Mouton de Gruyter, 2003.

Polinsky M. The Krymchaks: History and Texts. *Ural-Altaiic Yearbook*. 63. 1991. Pp. 123–154.

Seifart F., Evans N., Hammarström H., Levinson S. C. Language documentation twenty-five years on. *Language* 94 (4). 2018. Pp. e324–e345.