

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”**

1. Уточнение списочного состава языков России

А. А. Сюрюн, Т. И. Давидюк, А. П. Евстигнеева

Оглавление

1. Введение	3
2. Список языков России: предыдущие работы и исследования	4
3. Наш подход	8
3.1. Необходимость в регулярном обновлении списка.....	8
3.2. Учет целей ревитализации	9
3.3. Методика работы.....	12
4. Проблема “язык или диалект”	13
4.1. Краткая история противопоставления языка и диалекта в лингвистике ...	14
4.2. Критерии разделения языка и диалекта	15
4.2.1. Структурно-лингвистический критерий.....	16
4.2.2. Взаимопонятность	18
4.2.3. Социо-этно-политические критерии	20
4.2.4. Критерии, применяемые в международной организации по стандартизации ISO	21
4.2. Проблемы применения критериев разделения языка и диалекта	22
4.3. Проблема “язык или диалект” в рамках языков России	23
4.3.1. Абхазо-адыгские языки	25
4.3.2. Индоевропейские языки	26
4.3.3. Корейский язык.....	29
4.3.4. Нахско-дагестанские языки	30
4.3.5. Нивхский язык	35
4.3.6. Тунгусо-маньчжурские языки	36
4.3.7. Тюркские языки	36
4.3.8. Уральские языки	39
4.3.9. Чукотско-камчатские языки	46
4.3.10. Эскимосско-алеутские языки	47
4.3.11. Юкагирский язык.....	47
5. Экстратерриториальные языки России.....	48
6. Заснувшие языки	53
7. Другие проблемы	57
8. Перспективы дальнейшей работы	59
9. Заключение	59
10. Список использованной литературы.....	61
11. Приложения	74

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 составить актуальный список языков России. Эта задача является одним из ключевых вопросов подготовки Программы по сохранению и возрождению языков народов России, разрабатываемой в Институте языкознания РАН.

Для выполнения этого поручения было проведено всестороннее исследование соответствующей проблематики как в теоретическом, так и в прикладном аспектах. Были проанализированы релевантные социологические, этнологические, политические, географические и лингвистические параметры, определяющие таксономический статус идиомовов, распространенных на территории России. Кроме того, была исследована внутренняя структура каждого языка, обычно включающего по крайней мере несколько локальных идиомов. Отдельное внимание было уделено проблеме так называемых “экстратерриториальных” и “заснувших” языков.

Итоговый документ представляет собой описание теоретической базы и хода проведенного исследования. В нем последовательно рассматривается история вопроса и различные проблемы, связанные с составлением списка языков России, которые были характерны для предшествующих исследований. Далее в документе обосновывается используемый методологический подход, ориентированный на поддержку и сохранение языков, и подробно анализируются параметры, релевантные для определения статуса языка. На основе выработанных критериев предлагаются и подробно обосновываются решения для каждого языка со спорным статусом. В заключительной части документа обсуждаются перспективы дальнейшей работы и подводятся итоги проведенного анализа.

Составленный в ходе аналитической работы список языков послужил базой для проведения исследований по социолингвистическому состоянию языков России (см. Итоговый документ 2), методов поддержки языков (см. Итоговый документ 3), а также задокументированности языков (см. Итоговый документ 7).

2. Список языков России: предыдущие работы и исследования

В первом приближении список языков, распространенных на территории России, может быть сформирован на основании данных переписей населения. В нашей стране процедура переписи имеет довольно давнюю традицию. Первая всеобщая перепись населения состоялась в Российской империи в 1897 году. Следующая перепись, проведенная в 1920 году, не считается всеобщей из-за территориальной нерепрезентативности. В советский период переписи населения проходили приблизительно каждые десять лет (с перерывом в 1940-х). Такая же периодичность соблюдается и в наши дни: первая в истории современной России перепись прошла в 2002 году, вторая — в 2010 году, а третья, 2021 года, находится сейчас на завершающем этапе.

Переписные листы содержат значительное число вопросов, касающихся демографических, экономических и социальных данных населения страны. Релевантной для составления списка языков является информация об этнической самоидентификации, родном языке¹ и языках, которыми владеет опрашиваемый.

Однако данные, полученные таким способом, к сожалению, не лишены недостатков. Так, в результатах переписи 2002 года объединены данные по численности носителей иврита и идиша, хотя эти языки относятся к разным языковым семьям. В списке также фигурирует единый мордовский язык, который общепризнанно делится на эрзянский и мокшанский языки. Похожая ситуация встречается и в переписи 2010 года: в данных встречаются юитский и эскимосский языки, которые, как известно, подразделяются на несколько языков. В той же переписи встречается адыгский язык, который является общим названием для адыгейского и кабардино-черкесского языков.

Кроме того, переписная статистика может существенно отличаться от истинного числа владеющих некоторым языком. Во-первых, переписчиков, в отличие от лингвистов, не интересует уровень владения языком опрашиваемых². Например, несоответствие между переписной и реальной статистикой отмечается в статье БРЭ про водский язык: “По дан-ным пе-ре-пи-си 2010, чис-ло вла-дею-щих В. я. 68 чел., од-на-ко оно за-вы-ше-но – по оцен-кам лин-гвис-тов, В. я. вла-де-ют не бо-лее двух де-сят-ков чел.” [Агранат 2018]. Похожая ситуация наблюдается также и с другими языками. Так,

¹ Вопрос о родном языке отсутствовал в опроснике переписи 2002 года, но был восстановлен в опросниках переписи 2010 года.

² Тем не менее, в переписных листах для ВПН-2021 необходимо помечать галочкой языки, которые используются опрашиваемым в повседневной жизни. Это может дать некоторое представление о текущем уровне владения человека некоторым языком.

в переписи 2002 года 378 тофа указали, что на владеют тофаларским языком, однако в работах лингвистов Д. Харрисона и Г. Андерсона, проводивших в это же время исследования в местах компактного проживания этой этнической группы, говорится, что полноценных активных носителей тофаларского насчитывалось всего 20 человек, из них 14 в Алыгджере и Нерхе [Harrison 2003: 54, Harrison, Anderson 2008: 248]. Расхождение между данными переписей и реальным положением дел обсуждается также в [Намруева 2010] на материале калмыцкого языка. Во-вторых, искажение фактических данных может заключаться в ошибках переписчиков и обработчиках собранных данных, см., в частности, [Беликов 2020; Калимуллин 2003].

Еще одним источником информации для составления списка языков являются социолингвистические энциклопедии. В СССР такой крупнейшей энциклопедией являлась энциклопедия «Языки народов СССР», изданная в 1960-х гг. в пяти томах [Виноградов (отв. ред.) 1966а, 1966б, 1966в, 1967, 1968]. В данной энциклопедии дается описание 122 языков, а также пяти языковых семей и девятнадцати ветвей и групп. Перечень народов СССР приводится в соответствии с данными переписи 1959 г., однако в энциклопедии описываются языки не всех народов. Так, в ней не представлены немецкий, польский, болгарский, корейский, греческий, венгерский, румынский³, финский, китайский, персидский, словацкий, чешский языки, так как основная масса носителей этих языков живет за рубежом. Помимо языков народов, перечисленных в списке национального состава СССР в 1959 г., в энциклопедию включены малочисленные языки (такие как андийский, ботлихский, годоберинский, саамский, ительменский и др.), носители которых, как сообщается, причислили себя к представителям крупных наций, среди которых они живут. Авторы отмечают, что принципы включения или невключения в энциклопедию того или иного языка не всегда выдержаны: так, в список оказывается включен саамский язык, хотя его носители в большинстве своем проживают за рубежом [Виноградов (отв. ред.) 1966а: 12-13].

Социолингвистическая энциклопедия в двух томах «Письменные языки мира: Российская Федерация» [Солнцев, Михальченко (отв. ред.) 2000, 2003] является частью многотомного международного труда «Письменные языки мира», создававшегося по единому образцу учеными из университета им. Лавалья (Канада) и их коллегами из разных регионов мира. Первый том российского издания посвящен описанию языков с

³ При этом в энциклопедии имеется статья о молдавском языке, который многие лингвисты в настоящее время считают тем же самым языком, что и румынский.

непрерывной письменной традицией, а второй – языков с прерванной традицией, а также новописьменных и бесписьменных языков. Планировался также третий том с описанием языков этнических групп, народы-носители которых имеют государственность за пределами России, но он так и не был выпущен. Всего авторами выделяется 180 языков России, однако в вышедших двух томах представлено и описано только 84 из них. Описание принципов, по которым языки включались в энциклопедию, отсутствует, как и перечень всех 180 языков.

В трехтомной энциклопедии “Языки Российской Федерации и соседних государств” [Ярцева и др. (ред.) 2001; Виноградов и др. (ред.) 2001; 2005] содержатся статьи по 150 языкам. Также дополнительно есть 30 статей, посвященных языковым группам и семьям. Как сообщается в предисловии к третьему тому, в данной энциклопедии впервые описываются диалекты некоторых зарубежных литературных языков, распространенных в России и странах Ближнего Зарубежья (немецкие диалекты Поволжья, греческие диалекты на Украине, периферийный польский диалект в Литве, финский язык в Карелии и Ленинградской области). Критериев, по которым языки отбирались для статей энциклопедии, не приводится.

В статье Большой российской энциклопедии “Языки народов России” [Алексеев 2004] делается обзор языков тех народов, носители которых живут на территории России и не имеют государственных образований вне Российской Федерации. Сообщается, что таких языков около 100.

Последней опубликованной на данный момент социолингвистической энциклопедией на русском языке является “Язык и общество” [Михальченко (гл. ред.) 2016], которая продолжает исследование социолингвистических проблем, предпринятое в “Письменных языках мира”. Энциклопедия организована в виде словаря, и кроме статей, посвященных социолингвистическим портретам языков коренных народов России, в ней представлен понятийный аппарат социолингвистики. В сборнике присутствует информация по 93 языкам России, включая 30 языков коренных народов России и 63 языка малочисленных народов России.

Кроме того, ранее проводилась и целенаправленная работа по составлению списка языков России. В 2010 году по запросу Государственной Думы РФ Ю. Б. Коряковым был составлен список языков, наречий, диалектов и говоров Российской Федерации. В качестве самостоятельных языков в нем указана 121 единица. С 2019 года в Институте языкознания РАН ведется проект “Языки России” (руководитель проекта также Ю. Б. Коряков), целью которого является составление списка языков России по

четко выработанным критериям. В списке, разработанном в 2019 году, значится 153 языка. Языки включались в него на основании двух критериев — автохтонности и / или компактности проживания носителей языка на территории РФ. Согласно критерию автохтонности, языками России являются те языки, носители которых проживают на территории России более двух веков. Что касается второго критерия, то считается, что носители некоторого языка проживают компактно, если существуют поселения, в которых процент говорящих на этом языке составляет более 20% населения.

Некоторые сведения о списочном составе языков России можно почерпнуть и из зарубежных источников. Например, в ресурсе Ethnologue⁴ приводится 161 язык России⁵ [Eberhard, Simons, Fennig (eds.) 2021]. При этом три из них (бурятский, коми, марийский) являются так называемыми макроязыками (они подразумевают под собой несколько близкородственных языков, рассматривающихся в некоторых контекстах как один язык), поэтому они могут быть исключены при подсчете языков России. Стоит отметить, что в список включены не только родные языки народов России, но и выученные языки, такие как английский, испанский, итальянский, французский, норвежский и некоторые другие. Еще одним ресурсом является Glottolog⁶, в котором в качестве языков Российской Федерации фигурируют 160 языков⁷ [Hammarström et al. 2021]. Однако этот список вызывает много вопросов. Так, в него входят в том числе мертвые языки (например, маторский, печенежский и аринский языки). Представленность некоторых языков (например, ороchonского языка) на территории РФ также вызывает сомнения⁸. Кроме того, статус некоторых идиомов как отдельных языков представляется неоднозначным: так, в качестве отдельных языков выделены вах-васюганский хантыйский и сургутский хантыйский, хотя в мировой лингвистике они считаются диалектами хантыйского языка (или кантыкского (восточнохантыйского), если выделять севернохантыйский в самостоятельный язык).

Из представленных выше данных видно, что количество языков сильно разнится от работы к работе. В первую очередь на это влияют критерии, которые лежат в основе составления списка и которые выбирают сами исследователи исходя из своих задач. К

⁴ Ресурс доступен по ссылке <https://www.ethnologue.com/>.

⁵ Список доступен по ссылке <https://www.ethnologue.com/country/RU/languages>.

⁶ Ресурс доступен по ссылке <https://glottolog.org/>.

⁷ Список доступен по ссылке <https://glottolog.org/glottolog/language.map.html?country=RU#3/59.74/100.22>. Бурятский, коми и марийский языки, как и в предыдущем списке, являются макроязыками.

⁸ Носители ороchonского языка проживают в Китае, хотя до XVII в. проживали к северу от р. Амур. Вероятно, поэтому данный язык был включен в список, представленный в ресурсе Glottolog.

сожалению, в большинстве представленных работ такие критерии эксплицитно не указаны. Четкие критерии представлены только в работе проекта “Языки России”, возглавляемого Ю. Б. Коряковым. Также некоторые обоснования для включения языков в список приводятся в энциклопедии “Языках народов СССР”, однако сами авторы отмечают, что критерии отбора языков не всегда были выдержаны (см. выше). Кроме того, список языков России может меняться со временем (см. Раздел 3). Именно поэтому нашей задачей являлось составление актуального списка языков РФ с указанием критериев их включения исходя из конкретных задач нашего проекта (работа по сохранению, возрождению и документации языков России).

3. Наш подход

3.1. Необходимость в регулярном обновлении списка

Список языков какой-либо страны может меняться со временем, поэтому существующие перечни необходимо регулярно корректировать с учетом новых демографических данных. Так, необходимо учитывать миграцию: какие-то народы могут появляться на территории страны (например, в настоящее время больше половины удин проживает на территории России [Коряков 2020], хотя их историческим местом проживания является Азербайджан; казахи в 1930-х гг. переселились в Саратовскую область из Казахстана во время массового голода), какие-то, наоборот, уезжать (например, переезд по программе репатриации немцев из Поволжья и Сибири в Германию, а евреев — в Израиль). Народы могут ассимилироваться, а языки исчезать (ср. исчезновение беринговского диалекта алеутского со смертью последнего носителя в 2021 г.). Также могут изменяться границы государства: так, с присоединением Крыма к РФ в 2014 г. в состав языков России вошёл крымскотатарский, являющийся на этой территории государственным языком. Языки могут проникать в страну и другими способами: так, понтийцы, живущие в России, выучивают новогреческий язык в школе [Д. Ю. Зубалов, л. с.], из чего можно заключить, что в России постепенно растет число носителей этого языка. Отследить эти изменения во многом помогает Всероссийская перепись населения, поэтому результаты последней переписи, которая сейчас находится на завершающем этапе, могут повлиять и на состав предлагаемого нами предварительного списка языков России.

Кроме того, языки сами по себе подвержены изменениям. Так, 250 лет назад у коми-язьвинского и коми-пермяцкого языков были незначительные диалектные

различия, тогда как сейчас коми-язьвинский и коми-пермяцкий идиомы разошлись настолько, что их следует считать двумя разными языками [Норманская 2020]. Языки могут меняться и под действием окружающих их идиомов: к примеру, в XVII-XIX вв. часть шорцев переселилась в Хакасию, в результате чего их язык сблизился с хакасским языком и на данный момент считается диалектом хакасского [Межекова 1973]. Наши представления о составе языков также могут меняться с получением новой лингвистической информации. Например, появление стословных списков по все большему числу даргинских идиомов может приводить к изменению количества языков, которые выделялись из традиционного даргинского (ср. [Koryakov 2002, Коряков 2012, 2021; Коряков и др. 2020]). Другой пример касается мансийского языка, который в списке проекта “Языки России” разделяется на четыре разных языка — северномансийский, южномансийский, восточномансийский и западномансийский (последние три входят в список вымерших языков). Это решение основано на лексикостатистических подсчетах ([Zhivlov 2014; Коряков 2017]), выполненных по данным словаря Б. Мункачи, который был подготовлен на основе записей конца XIX в. [Munkácsi, Kálmán 1986]. Однако в более новой работе [Федотова 2021], учитывающей полевые данные 2010-х гг., постулируется единство мансийского языка. Таким образом, можно сделать вывод, что за прошедшее столетие мансийские идиомы значительно сблизились.

3.2. Учет целей ревитализации

Кроме того, что наш проект направлен на составление актуального списка языков России с учетом последних демографических и лингвистических данных, мы считаем важным смотреть на эту задачу через призму поддержки и возрождения (ревитализации) языков, которые в этом нуждаются. Дело в том, что если у некоторого идиома имеются какие-либо основания считаться не языком, а диалектом (о проблеме “язык или диалект” см. раздел 4), то наличие угрожаемого статуса может являться дополнительным фактором для признания этого идиома самостоятельным языком, так как это, например, открывает возможность издания учебных материалов. В последние десятилетия к теме ревитализации и поддержки малочисленных идиомов привлекается все больше внимания.

Одной из крупнейших работ, посвященных малочисленным языкам, является “Атлас языков, находящихся под угрозой исчезновения” (англ. Atlas of the World’s Languages in Danger) ЮНЕСКО. Первое издание Атласа было выпущено в 1996 году

[Wurm (ed.) 1996]. Обновленное издание Атласа увидело свет в 2001 году [Wurm (ed.) 2001]. Оно включало в себя описание 800 языков. С 2005 года ведется регулярно обновляемая онлайн-версия Атласа⁹ [Moseley (ed.) 2010]. Последняя печатная версия Атласа выпущена в 2010 году; она содержит около 2500 языков (из которых 230 исчезли начиная с 1950 года). Для каждого языка в Атласе указывается его название, уровень жизнеспособности (вычисляемый по специально разработанной ЮНЕСКО шкале, см. ниже) и страна или страны, где на нём говорят. Онлайн-издание предоставляет дополнительную информацию о числе носителей каждого языка, необходимых мерах по его сохранению, проектах, связанных с этим языком, об источниках информации, соответствующих кодах ISO 639-3 и географических данных.

Сведения об исчезающих языках мира также содержатся в работе “Encyclopedia of The World’s Endangered Languages” [Moseley (ed.) 2007]. Во введении к книге подробно обсуждаются причины, по которым большая половина существующих в мире языков находятся под угрозой исчезновения. Небольшие статьи по имеющимся в энциклопедии языкам предоставляют информацию о территории распространения языков, количестве их носителей и социолингвистической ситуации. Составители каталога придерживаются мнения о более дробном делении языков, особенно если информация про некоторые исчезающие идиомы является скудной [Moseley (ed.) 2007: 10-11]¹⁰.

Еще одним зарубежным проектом, аккумулирующим информацию по языкам, находящимся под угрозой исчезновения, является Проект по исчезающим языкам мира¹¹ (англ. Endangered Languages Project). Информация о языках для этого проекта берется из каталога исчезающих языков (англ. The Catalogue of Endangered Languages, или ELCat), который, в свою очередь, является результатом совместной работы Гавайского университета в Маноа и Университета Восточного Мичигана [Catalogue of Endangered Languages 2021]. В каталоге представлены языки, на которых говорят в 180 странах из 12 регионов мира. Среди исчезающих языков, распространенных в России, в каталоге выделяется 112 идиомов.

Что касается российского опыта, то в начале 2000-х гг. была предпринята попытка инвентаризации языков, над которыми нависла угроза исчезновения. В

⁹ Онлайн-версия Атласа доступна по ссылке www.unesco.org/languages-atlas/en/atlasmap.html.

¹⁰ “We have necessarily favoured ‘splitting’ over ‘lumping’ rather than run the risk of omitting what might emerge as a unique disappearing language, for want of information”. [Moseley (ed.) 2007: 11].

¹¹ Сайт проекта доступен по ссылке: <https://www.endangeredlanguages.com/>.

энциклопедическом справочнике «Языки народов России. Красная книга» [Нерознак 2002] представлено 66 таких идиомов. Уже из названия видно, что ориентация в этой работе сделана на идиомы этнических групп: за основу был взят перечень коренных малочисленных народов РФ. Поэтому иногда в качестве языков выделяются идиомы, которые традиционно считаются диалектами: так, выделяемый в книге язык шапсугов является диалектом адыгейского языка, а язык бесермян – диалектом удмуртского.

Еще одним источником информации об исчезающих языках является сайт «Малые языки России»¹², разрабатывающийся на базе Лаборатории исследования и сохранения малых языков ИЯз РАН. Главной задачей данного ресурса является систематизация имеющихся сведений о структуре и функционировании миноритарных языков России и их локальных вариантов, а также регулярное пополнение этих сведений актуальной информацией (см. [Поливанов и др. 2020; Polivanov et al. 2020]). Статья по каждому языку представляет собой шаблон, содержащий следующие типы информации: генетическая принадлежность, территория распространения, языковые контакты, сферы функционирования языка, динамика развития языковой ситуации, структура языка, сведения об исследователях и исследовательских центрах, список основных публикаций и электронных ресурсов, информация о наличии административной и общественной поддержки. В перспективе предполагается выложить на сайт информацию о всех языках России (проект ориентируется на список языков России, составленный Ю. Б. Коряковым и его коллегами [Коряков и др. 2020]).

Говоря об исчезающих языках, нельзя обойти вопрос о том, каким образом можно оценить состояние и дальнейшие перспективы функционирования какого-то языка. В частности, для этих целей в конце прошлого века начали создаваться различные шкалы измерения витальности (жизнеспособности) языков. Многие из этих шкал используются в энциклопедиях и ресурсах, описанных выше. Так, ЮНЕСКО использует свою собственную одноименную шкалу, которая насчитывает 6 уровней и строится на основе 9 критериев, наиболее значимым из которых является межпоколенческая передача языка. Ресурс Ethnologue использует расширенную многоступенчатую шкалу межпоколенческой утраты языка (Expanded Graded Intergenerational Disruption Scale, или EGIDS). При построении этой шкалы учитываются разные параметры: с одной стороны, официальный статус и сферы использования языка, с другой — владение языком представителей разных поколений. Шкала каталога исчезающих языков (The Catalogue

¹² Сайт проекта доступен по ссылке: <https://minlang.iling-ran.ru/>.

of Endangered Languages, или ELCat) основана на Индексе уязвимости языков (Language Endangerment Index, LEI), который вычисляется исходя из 4 факторов: межпоколенческая передача (intergenerational transmission), абсолютное число носителей (absolute number of speakers), соотношение численности говорящих и этнической группы (speaker number trends), сферы использования (domains of use). Более подробный обзор существующих шкал витальности и применение шкалы ELCat к материалу конкретных миноритарных языков России см. в работе [Казакевич и др. 2021].

Таким образом, при включении того или иного идиома в список языков России нам важно обращать внимание на то, представлен ли он в одном из перечисленных источников как исчезающий, и какой статус витальности приписывается ему на различных ресурсах. Кроме того, необходимо учитывать мнение экспертов и сотрудничать в этом вопросе с группой по уточнению состояния языков России. Повышение статуса какого-либо спорного идиома до отдельного языка (при наличии достаточных оснований) в случае, если он находится под угрозой исчезновения, могло бы привлечь внимание лингвистов и общественности к его поддержке и ревитализации.

3.3. Методика работы

Нами были изучены существующие на данный момент списки языков России (см. раздел 2). На основании полученной информации был составлен максимально широкий перечень идиомов (содержащий 229 единиц), которые потенциально могли бы быть включены в список языков России.

Основную проблему для составления списка языков представляет вопрос о том, какие идиомы следует считать языками, а какие — диалектами. Кроме того, для составления списка языков России представляют трудность языки, основная территория распространения которых находится за пределами нашего государства, однако их носители проживают также и в России (например, эстонский, казахский, таджикский и другие языки). Третью проблему представляет учет языков, признанных на территории России вымершими. Дело в том, что благодаря работе языковых активистов и лингвистов вымершие языки могут возвращаться к жизни, как, например, произошло с мэнским языком в Великобритании. Именно поэтому далее мы предпочитаем использовать словосочетание “заснувшие языки”.

Кроме того, в ходе работы были обнаружены и другие проблемы, которые могут быть связаны с некоторыми другими аспектами задачи составления списка языков России (см. подробнее раздел 7).

Параллельно с составлением максимально широкого списка “кандидатов” в языки России нами была изучена литература по возможным критериям составления списка языков. Для анализа проблемных случаев были отобраны только те критерии, которые, как нам кажется, являются наиболее релевантными в свете целей и задач Программы. Критерии были сгруппированы по трем проблемным случаям, указанным выше (проблема “язык или диалект”; проблема языков, распространенных преимущественно не в России; проблема заснувших языков). Каждый идиом, попадающий в какую-либо проблемную группу, был проанализирован с точки зрения релевантных для проблемы критериев. Сбор информации по проблемным идиомам осуществлялся по различным источникам, включая личное общение с экспертами-лингвистами и языковыми активистами¹³. Затем на основании полученных сведений нами принималось решение о том, включать или не включать тот или иной идиом в предварительный список языков России. “Непроблемные” языки включались в указанный список автоматически.

4. Проблема “язык или диалект”

Для составления списка языков в целом (не только списка языков России) наиболее существенной является проблема “язык или диалект”, заключающаяся в том, следует ли считать тот или иной идиом отдельным языком или же диалектом какого-то языка. Решение этого вопроса оказывается существенным для множества задач — для подсчета статистики (например, сколько языков представлено в том или ином регионе), создания языковых карт, составления грамматик, справочников, энциклопедий, баз данных и т. п. Важно отличать язык от диалекта также и для задач сохранения и документации языкового разнообразия, так как, как правило, программы поддержки со стороны государства обычно направлены на языки, а не на диалекты.

¹³ Мы проводили интервью со следующими специалистами: Л. Ф. Абубакирова, Т. Б. Агранат, Ф. Г. Алексеев, Н. Б. Аралова, В. Б. Бакула, Е. Ю. Груздева, И. А. Грунтов, Н. Р. Добрушина, А. В. Дыбо, В. А. Ельяшевич, М. М. Зимин, Р. В. Ишмухамбетов, С. Казама, Е. Н. Кондратьева, Ю. Б. Коряков, Н. В. Кропотова, Ю. А. Ландер, Е. А. Либерт, А. В. Мазарчук, Т. А. Майсак, Ю. Митров-Шимко, А. Д. Мусаев, М. З. Муслимов, Р. О. Муталов, И. Нагасаки, Ю. Нагаяма, Ч. Оно, С. А. Оскольская, Д. Т. Пиркова, С. С. Сай, Д. С. Сатанова, Л. С. Селендили, А. Слобожанин, В. В. Совкина, Я. Г. Тестелец, Д. М. Токмашев, Н. С. Уртегешев, А. Ю. Фильченко, М. С. Федина, В. С. Харитонов, Ю. Янхунен. Приносим им свою искреннюю благодарность за помощь.

4.1. Краткая история противопоставления языка и диалекта в лингвистике

Проблема “язык или диалект” в ее сегодняшнем виде возникает лишь в начале XIII века: до этого, начиная с Античности, понятия “язык” и “диалект” отсылали к письменной и устной речи соответственно (см., в частности, [Kamusella 2016; Van Rooy 2020]). Английский философ XIII века Роджер Бэкон противопоставил термины *lingua* и *idioma*, предвосхитив современное понимание различия между языком и диалектом. В эпоху Возрождения в трудах нидерландского гуманиста Йоханнеса Бекануса появляются рассуждения о взаимопонятности диалектов и языков; позже наличие или отсутствие взаимопонятности станет одним из современных критериев определения границы между языком и диалектом. Определение диалекта как разновидности некоторого языка дает в своих трудах швейцарский ученый XVI века Конрад Геснер.

В XVI веке набирают обороты проекты по стандартизации ряда европейских языков, связанные с распространением книгопечатания. В связи с этим границы между различными языками становятся более жесткими. В это же время становится популярной идея о том, что диалекты характерны для провинции, тогда как на “правильном” языке говорят жители столицы и ее окрестностей. Это приводит к тому, что диалекты начинают восприниматься как некоторое отклонение от стандартного языка.

К XVII веку складывается мнение, что разница между диалектами должна быть незначительной, в отличие от отдельных языков, которые существенно отличаются друг от друга. В конце XVII века появляется диалектология — раздел языкознания, изучающий территориальные диалекты и диалектные различия в их современном состоянии и истории. В XVII веке диалектология особенно активно развивается в Германии, где был осуществлен сбор материалов по говорам немецкого языка.

В XVIII веке начинают появляться подходы, пытающиеся инструментально отделить языки от диалектов. Так, немецкий исследователь Иоганн Кристоф Гаттерер предвосхитил появление лексикостатистики, попытавшись измерить различия между языками и диалектами с опорой на лексические совпадения в некотором списке базовой лексики.

В XIX веке благодаря трудам немецкого лингвиста Августа Шлейхера в языкознании появляется концепция генеалогического древа, заимствованная из области биологии. Для такого графического представления родственных языков оказывается

важным противопоставление языка и диалекта; язык приравнивается к биологическому виду, а диалект — к подвиду. Концепция генеалогического древа, как и четкое отделение языков и диалектов друг от друга, критиковались Гуго Шухардтом, Жюлем Жильероном и итальянской школой неолингвистики.

В 1950-е гг. получают развитие различные методы измерения взаимопонятности между идиомами, а также лексикостатистический подход.

В генеративной грамматике, получившей распространение в конце 1950-х гг., разница между языком и диалектом основывается на их структурном различии: диалекты имеют незначительные отличия в правилах, параметрах и ограничениях, проявляющихся в их структуре, в то время как языки характеризуются большими расхождениями в структуре. В социолингвистике, начавшей активно развиваться в 1950-1970-е гг., делается акцент на том, что различия между идиомами формируются в результате не только лингвистических, но и социополитических факторов. Так, в 1960-е гг. набирает популярность разработанная немецкими социолингвистами аусбау-парадигма, которая предполагает, что существует два независимых набора критериев, которые можно использовать для различения языка и диалекта: этносоциальные и структурные критерии (см. подробнее [Kloss 1967]). В работе [Haugen 1966a] также говорится, что существует два подхода к пониманию терминов “язык” и “диалект” — структурный и функциональный.

В настоящее время наблюдается тенденция отказа от жесткой дихотомии языка и диалекта. Так, в изданной Дэвидом Долби в 2000 году базе данных языков мира “Реестр Лингвосферы” применяется оригинальная трехуровневая классификация: внешний язык — внутренний язык — диалект (см. подробнее [Dalby 2000]).

4.2. Критерии разделения языка и диалекта

В [Перехвальская 2012] выделяются следующие критерии разделения языка и диалекта: 1) структурно-лингвистический критерий; 2) критерий взаимопонятности идиомов; 3) социолингвистические критерии.

В [Коряков 2017] структурно-лингвистический критерий и критерий взаимопонятности идиомов оказываются отнесены к внутренним критериям, а социо-этно-политические критерии — к внешним. Внешние критерии связаны не со структурными свойствами языка, а с его восприятием людьми. Внутренние критерии опираются на фактические различия между идиомами и являются независимыми от

мнения людей и других внешних факторов. Однако, как будет показано ниже, на взаимопонятность идиомов могут влиять экстралингвистические факторы, поэтому, как нам кажется, взаимопонятность не следует относить к внутренним критериям.

4.2.1. Структурно-лингвистический критерий

При помощи структурно-лингвистического критерия исследователи пытаются найти то лингвистическое расстояние (англ. *linguistic distance*, *linguistic difference*)¹⁴, при превышении которого отличающиеся друг от друга идиомы можно считать двумя отдельными языками. В целом для измерения лингвистического расстояния между двумя или несколькими идиомами было предложено множество методов: подсчет n-грамм [Singh, Surana 2007], сравнение частеречных тегов [Nerbonne, Wiersma 2006], деревьев зависимостей [Homola, Kuboň 2006], параллельных корпусов [Wälchli, von Wandelfels 2013], двуязычных словарей [Sejane, Eger 2013] и др. Довольно распространенным является метод сравнения грамматических свойств языков, применение которого, однако, связано с некоторыми трудностями. Во-первых, имеется необходимость определить набор сравниваемых свойств. Во-вторых, некоторые характеристики грамматической структуры языка могут быть связаны между собой (ср. теорию универсалий Джозефа Гринберга [Greenberg (ed.) 1963]), из-за чего одни грамматические свойства могут иметь больший “вес”, чем другие, см. работу [Hammarström, O’Connor 2013] о попытке учесть взаимозависимость признаков при измерении лингвистического расстояния. Для подсчета лингвистического расстояния в качестве исходных данных также часто используют списки Сводеша [СИГТЯ 2002: 713-737; Старостин 2012]. Сам М. Сводеш в рамках предложенного им метода глоттохронологии полагал, что одним и тем же языком следует считать идиомы, разошедшиеся менее чем 500 лет назад [цит. по: Яхонтов 1980].

Тем не менее, далеко не все перечисленные методы измерения лингвистического расстояния были применены к проблеме “язык или диалект”.

Методика разграничения языка и диалекта на основании процентных совпадений по спискам Сводеша предлагалась многими исследователями, см., в частности, [Duyn 1965; Kaufman 1994]. Четкое обоснование процентного порога, при нарушении которого идиомы могут считаться отдельными языками, содержится в работе [Коряков 2017]. Ю.

¹⁴ Под измерением лингвистического расстояния обычно понимаются различные метрики, показывающие, насколько различаются те или иные языки или диалекты.

Б. Коряковым были взяты процентные совпадения в 100-словном списке Сводеша между 220 парами близкородственных языков. Затем выбранные пары были разделены по пятибалльной шкале взаимопонятности (сведения о взаимопонятности были почерпнуты из языковых описаний). Кроме того, выбранные пары идиомов были распределены по их таксономическому классу в современных лингвистических работах (т. е. считаются ли они чаще диалектами или отдельными языками). Было показано, что диалектами одного языка чаще всего считаются идиомы, процент лексических совпадений между которыми составляет более 92%, а разными языками — с процентом лексических совпадений меньше 89%. В качестве рабочей гипотезы был принят порог в 90-91%: если процентное совпадение между 100-словными списками двух идиомов больше или равно 91%, то эти идиомы являются взаимопонятными и могут считаться диалектами одного языка. Диапазон между 90% и 91% является переходной зоной, и в случае, если идиомы попадают в эту область, следует принимать во внимание дополнительные факторы.

Работа [Wichmann 2019], ставящая перед собой задачу нахождения того лингвистического расстояния, при превышении которого идиомы могут считаться отдельными языками, также задействует лексемы из списка Сводеша, однако для сравнения используются только подмножества из 40 наиболее устойчивых слов [Holman et al. 2008]. Списки для более чем 5000 идиомов были взяты из базы данных ASJP [Wichmann, Holman, Brown (eds.) 2018]. Лексемы для разных языков в указанной базе данных записываются при помощи специальной транскрипции. Между лексемами различных пар идиомов было подсчитано нормализованное и деленное нормализованное расстояния Левенштейна¹⁵ (Levenstein distance normalized (LDN) и Levenstein Divided Normalized Distance (LDND)). Затем идиомы были классифицированы относительно их таксономического статуса (язык или диалект) согласно системе стандартизации ISO 639-3. Было выявлено, что значения $LDN = 0,51$ и $LDND \approx 0,599$ могут быть использованы в качестве пороговых значений для отделения пар диалектов от пар языков.

¹⁵ Расстояние Левенштейна определяется как наименьшее количество операций редактирования (вставки, удаления и замены), необходимых для получения одной текстовой строки из другой. Для вычисления нормализованного расстояния Левенштейна полученное расстояние необходимо поделить на длину большей строки. При вычислении деленного нормализованного расстояния Левенштейна нормализованное расстояние надо поделить на специальный коэффициент поправки.

4.2.2. Взаимопонятность

Взаимопонятность является одним из самых распространенных критериев разделения языка и диалекта. Степень взаимопонятности может зависеть не только от лингвистического расстояния между идиомами, но и от таких экстралингвистических факторов, как предварительное знакомство с языком собеседника и отношение слушающего к этому языку. Поэтому иногда выделяется два вида взаимопонятности — выученная взаимопонятность (*acquired intelligibility*; когда человеку понятен некоторый не родной ему идиом в силу того, что этот идиом в какой-либо степени знаком ему) и внутренняя взаимопонятность (*inherent intelligibility*; когда язык понятен носителю другого языка в силу лингвистической близости этих языков). За последние десятилетия появилось множество методов количественного измерения взаимопонятности. Критерий взаимопонятности является ведущим в современных базах данных и каталогах языков мира. Так, в системе Glottolog некоторый идиом признается отдельным языком, если он является невзаимопонятным по отношению к другим идиомам.

Основные методы измерения взаимопонятности между идиомами излагаются в работе [Gooskens 2013]. Предлагающиеся в литературе методы можно разделить на два типа — исследование мнения слушающего (*opinion testing*) и функциональное тестирование (*functional testing*).

Первый тип измерения взаимопонятности заключается в том, что носители одного идиома оценивают по некоторой шкале, насколько хорошо или плохо они понимают носителей другого идиома. Методы по исследованию мнения слушающего делятся на два вида по типу предлагаемых испытуемому вопросов. Один метод не предполагает предъявления испытуемому каких-либо образцов речи на другом идиоме: у человека интересуются, насколько при гипотетической встрече с носителем другого идиома ему понятна речь другого и, наоборот, насколько носителю другого идиома понятна речь опрашиваемого. Подобный метод был опробован на материале трех скандинавских языков — датского, норвежского и шведского — в работе [Haugen 1966b]. Второй вид измерения взаимопонятности, опирающийся на мнение слушающего, задействует образцы речи на неродном для слушающего идиоме. Испытуемый должен оценить, насколько ему было понятно то, что он прослушал; см. о применении подобного метода на материале нидерландских [van Hout, Münsterman 1981], норвежских [Gooskens, Heeringa 2004] и китайских диалектов [Tang, Van Heuven 2007].

Второй тип методов изучения взаимопонятности — функциональное тестирование (functional testing) — направлен на измерение не разборчивости речи в целом, а того, насколько слушающий реально понимает ее содержание. Ш. Госкенс выделяет следующие методики:

1. Вопросы по содержанию. Испытуемые должны отвечать на вопросы, касающиеся содержания прослушанного текста. Понятность идиома носителю другого идиома измеряется с опорой на количество правильных ответов, данных испытуемым.
2. Перевод. Испытуемые должны перевести на свой язык прослушанный текст. Чем больше слов человек переведет правильно, тем более для него, как гласит данный метод, понятна речь на другом идиоме. Эта методика была применена, в частности, к материалу скандинавских языков [Gooskens, Beijering, Heeringa 2008]. Одной из разновидностей теста на перевод является SPIN-тест (Speech Perception in Noise), в котором испытуемому необходимо перевести только последнее слово в предложении. SPIN-тест позволяет, в частности, минимизировать эффекты разного объема рабочей памяти испытуемых.
3. Пересказ текста. Метод RTT (Recorded text testing), подробно описанный в [Casad 1974], имеет две разновидности, в одной из которых испытуемый должен максимально точно пересказать текст [Kluge 2010]. Другой разновидностью метода RRT являются вопросы по содержанию текста.
4. Измерение времени реакции. Например, при использовании метода вопросов по содержанию текста можно измерять время реакции: чем дольше испытуемый думает над ответом, тем менее ему понятен услышанный текст. Измерение времени реакции в качестве дополнительного фактора было применено, в частности, в исследовании [James et al. 1994], где испытуемые должны были оценивать истинность или ложность услышанных ими предложений.

Расчет пороговых значений степени взаимопонятности, при нарушении которых два идиома могут считаться отдельными языками, был произведен только для метода RTT, см. [Blair 1990; Starr, Stalder 1990]. Стоит сказать, что целенаправленные исследования взаимопонятности на материале языков России практически не проводились (исключением являются некоторые даргинские идиомы, см. раздел 4.4.4.2). Как правило, в существующих грамматических описаниях содержатся лишь оценочные суждения лингвистов о том, как носители различных диалектов и близкородственных языков понимают друг друга. Отметим, что иногда лингвисты могут расходиться в своих

суждениях, см., например, противоречивые сведения о взаимопонятности между двумя энецкими идиомами, приводимые в [Коряков 2018a].

4.2.3. Социо-этно-политические критерии

Социо-этно-политические критерии не связаны со структурными свойствами языка. Перечислим некоторые из них:

1. Наличие общей литературной нормы: если у идиома есть собственная литературная норма, то он может считаться отдельным языком.
2. Функциональная полноценность: если идиом используется ограниченно (например, только в ситуациях бытового общения), то, как правило, он считается диалектом другого языка.
3. Этническое самосознание носителей: если носители какого-то идиома считают себя отдельной этнической группой, то их идиом можно определить как отдельный язык.
4. Мнение носителей относительно своего идиома: идиом считается самостоятельным языком, если его носители считают его таковым.
5. Официальный статус идиома: самостоятельным языком считается тот идиом, для которого в каком-либо официальном документе государства постулируется, что он является языком. Так, статус румынского и молдавского языков (которые, по мнению многих лингвистов, являются двумя названиями одного и того же языка) указывается в законодательстве Румынии и Молдавии соответственно и тем самым считаются с этой точки зрения разными языками.
6. Политический фактор: некоторый идиом в одном государстве может считаться диалектом, а в другом — отдельным языком (например, русинский во многих странах считается самостоятельным языком, тогда как в Украине — диалектом украинского языка); два лингвистически и функционально самостоятельных идиома, имеющих в стране, могут считаться диалектами одного языка (как, например, путунхуа и кантонский в Китае).

Перечисленные критерии часто оказываются более весомыми, чем лингвистические критерии и критерий взаимопонятности. Так, как указывается в [Campbell, Belew (eds.) 2018: 138-139], ведущим фактором при определении языков Европы является этническое единство носителей некоторого языка. Например, цаконский обычно рассматривается как диалект греческого языка, хотя у него отсутствует взаимопонятность с современным греческим языком, а процент

лексических совпадений в базовой лексике составляет 70%. С другой стороны, шведский и норвежский языки считаются отдельными языками, так как их носители осознают себя как отдельные этнические группы, хотя лингвистически эту пару идиомов можно считать одним языком.

Кроме того, необходимо обращать внимание на территориальную разобщенность носителей идиома. Наличие труднопреодолимых природных границ или больших расстояний может явиться предпосылкой для выделения в диалектном континууме лингвистически разошедшихся идиомов.

Дополнительным фактором для возникновения проблемы “язык или диалект” может стать различная религиозная принадлежность внутри этнической группы. Так, религиозная принадлежность оказывается одним из ключевых факторов различения отдельных идиомов в группе северо-восточных новоарамейских языков.

4.2.4. Критерии, применяемые в международной организации по стандартизации ISO

ISO 639 является международным стандартом кодов для представления названий языков. Стандарт состоит из пяти частей; для проблемы “язык или диалект” оказывается важной третья часть, ISO 639-3. Она задает трехбуквенный код, который присваивается всем языкам, имеющимся в мире, включая вымершие и древние языки. Эти коды являются открытыми списками и могут быть расширены и усовершенствованы. Работа по поддержанию стандартов ведется органами регистрации (Registration Authority); для ISO 639-3 таким органом регистрации является Летний институт лингвистики (Summer Institute of Linguistics; SIL International). В организацию можно подать запрос на присвоение некоторому идиому трехбуквенного кода, что по сути является признанием этого идиома отдельным языком. Базовыми критериями для различения языка и диалекта являются следующие факторы¹⁶:

1. Два родственных идиома обычно считаются разновидностями одного и того же языка, если носители каждого идиома понимают другой идиом (без необходимости его изучения) на функциональном уровне.
2. Если устная взаимопонятность между идиомами достаточно невелика, то наличие общей литературной нормы с центральным идиомом, понятным всем носителям, и общего этнолингвистического самосознания может быть весомым

¹⁶ См. <https://iso639-3.sil.org/about/scope#I>.

показателем того, что идиомы следует рассматривать как разновидности одного языка.

3. Если между идиомами имеется взаимопонятность, то наличие четко установленных и отличных друг от друга этнолингвистических идентичностей у носителей этих идиомов может быть убедительным показателем того, что эти идиомы следует рассматривать как отдельные языки. Этот параметр особенно важен в случае, когда языковые сообщества оказываются разделены политическими границами.

Таким образом, в системе стандартизации ISO ведущим критерием разделения языка и диалекта можно назвать этническое самосознание носителей.

4.2. Проблемы применения критериев разделения языка и диалекта

Универсального критерия, по которому можно было разделить языки так, чтобы это разделение не вызывало споров среди лингвистов и языковых сообществ, не существует.

Достаточно объективными в плане решения проблемы “язык или диалект” выглядят лексикостатистические методы, описанные нами в разделе 2.2.1, однако они нередко вступают в противоречие с экстралингвистическими критериями. Так, украинский и белорусский языки имеют 97% лексических совпадений, однако, тем не менее, считаются двумя самостоятельными языками из-за различного этнического самосознания носителей, а также в силу политических и исторических факторов. Напротив, множество групп значительно разошедшихся идиомов в настоящее время считаются одним языком. К примеру, процент лексических совпадений между основными городскими диалектами китайского языка не превышает 88% (см. [Wang 1994]), однако из-за экстралингвистических факторов они не считаются самостоятельными языками.

В [Hammarström 2008] было показано, что критерий взаимопонятности может быть использован для последовательного и непротиворечивого разбиения лингвистического разнообразия человечества на отдельные языки. Напомним, что на критерий взаимопонятности во многом опирается база данных языков мира Glottolog. Тем не менее, критерий взаимопонятности из-за его зависимости от различных экстралингвистических факторов критиковался многими исследователями вплоть до того, что от данного критерия предлагали отказаться, см., в частности, [Agard 1971]. Сложности вызывает явление несимметричной взаимопонятности. Так, во множестве

исследований было показано, что носители шведского языка хуже понимают носителей датского языка, чем те — шведский. Для объяснения этого явления предлагались как экстралингвистические (например, отсутствие позитивного отношения шведов к датскому языку [Delsing, Åkesonn 2005]), так и лингвистические факторы [Schüppert 2011]. Случаи несимметричной взаимопонятности показывают, что уровень взаимопонятности между одной и той же парой языков может быть различен. Кроме того, как и в случае со структурно-лингвистическим критерием, экстралингвистические критерии оказываются более весомыми, чем критерий взаимопонятности.

Наличие общей литературной нормы также не является универсальным критерием. Во-первых, далеко не все языки мира имеют письменность (о проблеме “язык или диалект” в применении к материалам бесписьменных языков см., в частности, [Эдельман 1980]). Во-вторых, существуют языки, в которых функционирует несколько систем письменности, ср. букмол и нюношк для норвежского языка.

Не все языки имеют какой-либо официальный статус и тем самым обладают неполной функциональностью. Следовательно, критерии наличия официального статуса и функциональной полноценности также не являются универсальными.

Зачастую язык приравнивается к нации или этнической группе, однако это не всегда так (подробнее о трудностях приравнивания языка к этнической группе см. [Перехвальская 1997]). Например, одна этническая группа может состоять из носителей нескольких языков. Так, швейцарцы говорят на четырех различных языках. Кроме того, разные этнические группы могут говорить на одном языке. Например, на испанском языке говорят и испанцы, и мексиканцы.

Не является универсальным критерием и мнение носителей идиома. Например, в исследовании [Adler 2021] показывается, что жители Германии имеют различное мнение по поводу того, можно ли считать отдельным языком нижненемецкий или нет.

4.3. Проблема “язык или диалект” в рамках языков России

Практически внутри каждой языковой семьи, распространенной в России, можно найти идиомы, чей таксономический статус (язык или диалект) вызывает вопросы и споры у исследователей. Мы проанализировали проблемные идиомы по следующим вопросам и параметрам:

1. Наличие индекса в ISO 639-3;
2. Географические параметры:
 - a. Ареал распространения носителей идиома;

- b. Проживают ли носители идиома отдельно от носителей того языка, к которому традиционно относится идиом?
- 3. Количество носителей (данные переписей, экспертные данные);
- 4. Наличие литературной нормы, ее единство;
- 5. В случае отсутствия литературной нормы — какой язык или вариант языка преподается в школе для носителей идиома?
- 6. Традиционность выделения идиома:
 - a. Имеются ли специалисты, считающие, что идиом следует объединить с другим идиомом (или другими идиомами) или разделить на несколько идиомов?
 - b. Выделяется ли идиом в качестве самостоятельного языка в российских и советских лингвистических энциклопедиях?
- 7. Лингвистические параметры:
 - a. Сведения о взаимопонятности с близкими идиомами;
 - b. Информация о процентном совпадении между близкими идиомами в базовой лексике;
 - c. Структурные различия между близкими идиомами;
- 8. Как называют свой идиом его носители?
- 9. Точка зрения носителей идиома (считают ли они свой идиом отдельным языком или диалектом другого языка?);
- 10. Считают ли носители идиома себя отдельной этнической группой?
- 11. Имеется ли внутри языкового сообщества разделение по исповедуемой религии?

Для того, чтобы на основании приведенных параметров можно было вывести итоговое решение, они должны образовывать некоторую иерархию. В основном мы ориентировались на лингвистические критерии и сведения о взаимопонятности, так как нарушенная взаимопонятность негативно влияет на статус витальности идиома. Так, если в школах учащиеся на уроках родного языка изучают язык, который существенно отличается от идиома, которым они владеют, то это может вызвать потерю интереса к занятиям и родному идиому в целом. Параметры наличия индекса в ISO 639-3 и традиционности выделения идиома в качестве самостоятельного языка не позволяют нам объединять близкие идиомы, которые принято считать отдельными языками (например, адыгские языки и коми языки). Остальные параметры мы рассматриваем как дополнительные аргументы за выделение того или иного идиома как самостоятельного языка.

Ниже мы рассмотрим проблемные идиомы по языковым семьям, укажем итоговое решение (являются ли они, с точки зрения выбранных нами критериев, отдельными языками или диалектами) и его краткое обоснование. Более подробный анализ проблемных идиомов см. в Таблице 2 в Приложении.

4.3.1. Абхазо-адыгские языки

В семье абхазо-адыгских языков можно выделить следующие проблемные идиомы: шапсугский диалект адыгейского языка, бесленеевский диалект кабардино-черкесского языка и ашхарский диалект абазинского языка.

Шапсуги выделяют себя как этническую группу, отличную от остальных адыгов, и, кроме того, входят в Единый перечень коренных малочисленных народов Российской Федерации. Однако мы не склонны считать шапсугский диалект самостоятельным языком. Шапсугский диалект отличается от других диалектов адыгейского языка несущественно. Кроме того, нет исследователей, которые считали бы шапсугский отдельным языком. Дополнительно отметим, что наблюдается значительная вариативность при использовании шапсугского различными поколениями говорящих [Lander 2010], что, возможно, впоследствии может повлиять на лингвистическое расстояние между шапсугским и другими диалектами адыгейского.

Бесленеевский диалект лингвистически является переходным между адыгейским и кабардино-черкесским языками [Коряков 2006: 24; Кумахов (ред.) 2006: 669]. Ашхарский диалект абазинского языка лингвистически довольно близок к абхазскому языку [Чирикба 2001: 2]. Тем не менее, указанные диалекты не обнаруживают каких-либо существенных отличий, которые бы значительно нарушали взаимопонятность между идиомами абхазо-адыгской семьи.

Таким образом, в предварительный список языков России оказываются включены абазинский, адыгейский и кабардино-черкесский языки (об абхазском языке см. раздел 5). Адыгейский и кабардино-черкесский языки иногда объединяются под названием “адыгский язык” (ввиду наличия общего самоназвания носителей этих языков — адыги). Мы не поддерживаем такое объединение, так как адыгейский и кабардино-черкесский языки обнаруживают значительные лингвистические отличия между собой и обладают отдельными литературными нормами; кроме того, носители этих языков в большинстве своем проживают в различных регионах РФ, то есть они оказываются разъединены административными границами.

В целом языки абхазо-адыгской семьи лингвистически довольно близки друг другу. В [Рogaвa 1967: 95] отмечается, что абхазский и абазинский в лингвистическом смысле можно считать диалектами одного языка. В настоящее время постулировать существование единого “абхазо-абазинского” языка представляется нерациональным из-за того, что абхазский и абазинский идиомы оказываются разделены государственными границами между Россией и Абхазией.

4.3.2. Индоевропейские языки

В осетинском языке выделяются иронский и дигорский диалекты, между которыми наблюдается наличие несимметричной взаимопонятности: из-за того, что литературный осетинский язык был основан преимущественно на иронском диалекте (примерно до конца XX века), носители дигорского языка достаточно хорошо понимают носителей иронского, но не наоборот [А. П. Выдрин, л. с.]. Затруднительность во взаимопонятности между носителями иронского и дигорского также отмечается в [Исаев, Камболов 2016]. Процент совпадений в общей базовой лексике между этими двумя диалектами составляет 92%, что превышает порог в 90-91% и не позволяет уверенно говорить о существенных различиях в базовой лексике [Коряков, Майсак 2001]. Лексикостатистические подсчеты по спискам для дигорского и иронского диалектов, содержащимся в базе данных ASJP, также не дают расхождения, достаточного для постулирования их отдельными языками ($LDN = 0,4576$). Литературная норма существует и для иронского, и для дигорского диалекта, хотя последняя была признана равноправной с первой только в конце XX века [Исаев, Камболов 2016]. Что касается этнического самоопределения, то и иронцы, и дигорцы считают себя осетинами. В связи с этим и с тем, что иронский и дигорский диалекты не обнаруживают существенных лингвистических отличий, мы вносим в предварительный список единый осетинский язык.

На территории России традиционно проживают носители амшенского и норахичеванского (донского) диалектов армянского языка. Амшенский армянский распространен также в Турции; условно амшенцев можно разделить на три группы — западные амшенцы (провинция Ризе, Турция), восточные амшенцы (провинция Артвин, Турция) и северные амшенцы (Абхазия и юг России). В [Vaughan 2007: 257] сообщается, что амшенский армянский разделяет многие лингвистические признаки восточноармянского языка, однако не является с ним взаимопонятным. Что касается лексикостатистических данных, то в [Коряков 2018б] было проведено предварительное

сравнение 100-словного списка для одного из северных амшенских диалектов со списком для восточноармянского литературного языка, которое показало, что процент лексических совпадений между ними равняется 80%, что существенно ниже условной границы в 91% между языком и диалектом. Таким образом, мы выделяем амшенский армянский как самостоятельный язык. Нор-нахичеванский (или донской) диалект относится к западноармянским диалектам [Мартirosян 2013]. В [Хастян 2001] затрагивается проблема этнического самоопределения различных групп армян в России; в работе было показано, что донские армяне довольно четко противопоставляют себя армянам, живущим в Армении. Данных о том, насколько донской диалект взаимопонятен с остальными западноармянскими диалектами, нам не удалось обнаружить. Так как в армяноязычном обществе бóльшую роль играет восточноармянский, а не западноармянский язык (последний, например, “Атласом языков мира, находящихся под угрозой исчезновения” был признан на территории Турции как безусловно находящийся под угрозой исчезновения (*definitely endangered*)), то мы вносим донской армянский как представителя западноармянских идиомов на территории России в предварительный список языков.

В России также проживают носители немецкого языка. Известно, что некоторые диалекты немецкого существенно отличаются от немецкого литературного языка. В последнее время немецкий рассматривают как два отдельных языка — верхненемецкий и нижненемецкий (подробную аргументацию о выделении нижненемецкого как самостоятельного языка см. в [Goltz, Kleene 2020]). На территории России встречаются носители как нижненемецкого, так и верхненемецкого [Брук, Смирнова 1994], поэтому оба эти языка включены нами в предварительный список языков России; подробнее о нижненемецком и верхненемецком языках в России см. также раздел 5.

Проживают в России и понтийские греки, чей язык также попадает под проблему “язык или диалект”. Понтийский имеет отдельный языковой код в ISO 639-3. В грамматике понтийского [Drettas 1997] автор дает позитивный ответ на вопрос, можно ли считать понтийский на основании структурно-лингвистических отличий от стандартного новогреческого самостоятельным языком. Поэтому мы вносим понтийский в предварительный список языков России; о нем, а также о новогреческом, см. также раздел 5.

В России проживает группа батумских курдов (курманч) [Тишков, Журавский, Казьмина (отв. ред.) 2008: 75]. Они являются носителями такого курдского идиома, как курманджи. Носителями курманджи также являются курды, проживающие в Республике

Адыгея, см. [Соколова (гл. ред.) 2015]. Курманджи в научной литературе считают либо языком, либо диалектом курдского. В ISO 639-3 курманджи присвоен отдельный языковой код. В [Hassanpour 1992] сообщается о наличии невзаимопонятности как минимум между курманджи и сорани (центральнокурдским диалектом). В недавней работе [Bourgeois-Gironde et al. 2021] на основании лексикостатистических подсчетов показывается, что лингвистическое расстояние между курманджи и сорани позволяет считать их самостоятельными языками, а не диалектами одного языка. Кроме того, курманджи имеет отдельную письменность. Мы выделяем курманджи как отдельный язык в нашем предварительном списке. Среди групп курдов, которые в большинстве своем исповедуют ислам, выделяются езиды, чьей религией является езидизм. Иногда язык езидов выделяется как самостоятельный язык, однако такая позиция критикуется из-за отсутствия проверенных доказательств [Ali 2020].

На территории России проживают также носители татских идиомов. Традиционно в татском языке выделяется два диалекта — северный и южный. Носители северного диалекта исповедуют иудаизм. Носители южного диалекта в большинстве своем являются мусульманами. Также к носителям южного диалекта относятся так называемые армяно-таты — армяне, перешедшие на татский язык; они исповедуют христианство. Как сообщается в [Грюнберг 1997], взаимопонятность между северным и южными диалектами, в том числе и в силу описанных конфессиональных факторов, затруднена. В результатах Всероссийской переписи 2002 года татский язык (язык мусульман и христиан; южный диалект татского по приведенной выше классификации) и язык горских евреев (северный диалект) фигурируют как два самостоятельных языка. В системе стандартизации ISO 639-3 им также присвоены два отдельных языковых кода. Разделены эти два идиома и в “Атласе языков мира, находящихся под угрозой исчезновения” ЮНЕСКО: язык горских евреев имеет статус “есть угроза исчезновения” (definitely endangered), а татский язык — “серьезная угроза исчезновения” (severely endangered). Мы в своем списке также разделяем горско-еврейский и татский языки; последнее наименование подразумевает под собой язык татов-мусульман и татов-христиан. Татский язык (язык татов-мусульман и татов-христиан) распространен в основном в Азербайджане, поэтому мы относим его также к экстратерриториальным языкам России. Тем не менее, в Дагестане, по-видимому, все же имеются компактные поселения носителей татского языка — в селах Джалган и Митаги-Казмаляр [Кусаева, Сатцаев, Таказов 2016].

В России также компактно проживают латгальцы (см. раздел 5) — носители латгальского языка. Латгальский имеет спорный статус в вопросе о том, является ли он языком или диалектом. Так, в Латвии латгальский признается вариантом латышского языка [Likums 1999]. Предварительные лексикостатистические подсчеты в базовой лексике указывают на сходство литературного латгальского и литературного латышского в 95% [Андронов 2008]. Тем не менее, множество лингвистов признают латгальский самостоятельным языком (см., к примеру, список таких исследователей в [Брейдак 2006]). В [Брейдак 2006] также приводятся сведения о нарушенной взаимопонятности между латышскими диалектами и латгальским. У латгальцев имеется четкая этническая самоидентификация и своя письменность с XVIII века. Кроме того, латгальскому присвоен отдельный языковой код в ISO 639-3. Латгальцы отличаются от латышей и по конфессиональной принадлежности: так, большинство латышей исповедуют лютеранство, тогда как латгальцы — католицизм. Таким образом, на основании приведенных фактов мы склонны считать латгальский отдельным языком.

4.3.3. Корейский язык

В России распространен корейский язык обоих государств: северокорейский и южнокорейский. Первое появление носителей северокорейских диалектов на территории современной России датируется 30 ноября 1863 г., когда двадцать корейских семей обратились к начальнику русского пограничного поста на юге Приморского края разрешить им поселиться в долине нынешней реки Виноградная в Хасанском районе Приморского края. Затем последовало активное переселение корейцев из северных провинций Кореи, переживавших массовый голод из-за сильных наводнений [Волынец 2020].

По результатам русско-японской войны южная часть острова Сахалин принадлежала Японии и называлась префектурой Карафуто. Сюда с 1905 года принудительно было переселено население с Корейского полуострова; большинство составляло жителей южной части Кореи.

Из-за накалившейся обстановки в отношениях с Японией, а также угрозы возможного нападения японцев на Дальний Восток власти СССР в 1937 году выслали с Дальнего Востока в Среднюю Азию, Казахстан и Узбекистан около 170 тысяч этнических корейцев [Волынец 2020]. В послевоенный период свыше 2,5 тысяч корейцев были отправлены из Казахстана и Узбекистана в специальные командировки в Северную Корею и Южный Сахалин [Ким 2019].

В настоящее время корейцев, проживающих на территории России, разделяют на материковых (добровольно переселившихся на Дальний Восток с 1860-х по 1920-е гг.) и сахалинских (принудительно переселенных Японией с 1905 по 1945 гг.).

Материковых корейцев называют корё-сарам, и они говорят на корё мар. В нашем списке идиом, на котором говорят корё-сарам является проблемным, так как это диалект северокорейского языка, который существует в основном в устной форме. Корё мар развился на основе субдиалекта хамгёнского диалекта югып, распространенного в шести северокорейских уездах - Пурён, Кёнхын, Онсон, Кёнвон, Косон и Хверён. Корё мар сильно отличается от сеульского и пхеньянского диалектов, лежащих в основе южнокорейского и северокорейского языков, соответственно. Корё мар долго находился в изоляции от основного языкового ареала, активно взаимодействовал с русским языком и языками Средней Азии, в результате чего фонетика, грамматика и лексика подверглись сильным изменениям. Идиом не преподается в школах, находится под угрозой исчезновения, так как активно используется лишь небольшим числом людей пожилого возраста. Большинство корё-сарам перешли на русский язык [Трофименко 2012].

Сахалинские корейцы в основной своей массе происходили из южных провинций. После присоединения Южного Сахалина к советской территории корейские дети получили возможность учиться в школах на корейском и русском языках, однако в мае 1963 года корейские школы были реорганизованы. Закрылись также Поронайское и Южно-Сахалинское корейские педагогические училища. Сейчас на корейских отделениях вузов России, в том числе и Сахалина, обучают сеульскому литературному языку [Ким 2019]. Во время интервью для новостного портала Sakhalin.info Светлана Ким сообщает, что молодежь, изучающая современный корейский язык, испытывает сложности в понимании варианта, на котором говорят пожилые корейцы, так как в настоящее время на Сахалине вместе с сахалинскими корейцами проживают и материковые корейцы, говорящие на корё мар, а также выходцы из Северной Кореи, которые были привезены в 1948 году в качестве рабочей силы [Герцен 2021].

Мы считаем, что имеются все основания выделять корё мар в качестве самостоятельного языка, таким образом, корё мар и идиом сахалинских корейцев (корейский язык) могут быть включены в список языков России.

4.3.4. Нахско-дагестанские языки

Идиомы, проблемные с точки зрения вопроса “язык или диалект”, есть среди аваро-андо-цезских, даргинских, лезгинских и нахских языков нахско-дагестанской семьи. В лакском языке, не относящемся к перечисленным группам языков, проблемных идиомов не обнаруживается, поэтому он был автоматически внесен в предварительный список языков России.

4.3.4.1. Аваро-андо-цезские языки

Для диалектов аварского языка пока не собраны списки Сводеша и, соответственно, нет достоверных данных лексикостатистики, но косвенные данные намекают на то, что это группа языков [Коряков 2020]. Мы оставляем решение этого вопроса для дальнейших исследований и вносим в список единый аварский язык.

Традиционно выделяется восемь андийских языков. Все эти языки довольно слабо изучены. Лингвистических данных мало, а основываться на мнении носителей не получается, потому что носители каждого села считают свой вариант языка особенным и не ощущают общности с вариантами языка соседних сел [Коряков 2020]. Непроблемными среди андийских языков являются: ботлихский, годоберинский, каратинский, тиндинский и багвалинский языки. Андийский язык делится на две группы говоров, верхнеандийскую и нижнеандийскую, которые в последнее время на основании данных лексикостатистики принято считать отдельными языками (см. [Koryakov, Verhees 2021]). Мы также придерживаемся этой точки зрения, выделяя верхнеандийский (или собственно андийский) и нижнеандийский (мунийско-кванхидатльский) языки. Ахвахский язык похожим образом традиционно делится на северный и южный диалекты, между ними имеются отличия фонетического, морфологического, лексического характера, взаимопонимание между диалектами затруднено [Алексеев 2016а]. Поэтому мы вслед за Ю. Б. Коряковым выделяем североахвахский и южноахвахский языки. При этом в южноахвахском есть три диалекта (ратлубский, цекобский, тлянубский), которые, возможно, в будущем могут быть признаны разными языками [Коряков 2020; Koryakov, Verhees 2021]. Проблемным случаем также является тукитинский диалект каратинского языка, который не является взаимопонятным с собственно каратинским. По данным лексикостатистики они значительно расходятся, поэтому мы выделяем тукитинский в самостоятельный язык. Еще одним своеобразным идиомом каратинского языка является анчихский говор, носители которого считают свой язык отдельным от каратинского. Однако вопрос лингвистической разницы и степени взаимопонятности анчихского и собственно каратинского требует дальнейших

исследований. Гигатлинский считается более архаичным диалектом чамалинского языка, взаимопонимание между носителями гигатлинского и второго чамалинского диалекта, гакваринского, затруднено [Алексеев 2016б]. В [Koryakov, Verhees 2021] высказывается предположение о выделении гигатлинского в самостоятельный язык, то же подтверждает наше личное общение с Н. Р. Добрушиной и Т. А. Майсаком. Мы поддерживаем это разделение.

Традиционно выделяется пять цезских языков. Среди них непроблемными являются: цезский, гинухский, бежтинский и гунзибский языки. Процент лексических совпадений между сагадинским и цезским 89%, взаимопонимания между ними нет [Коряков 2020], поэтому мы считаем сагадинский отдельным языком. Инхокваринский и хваршинский имеют процент совпадений 93% [Коряков 2020], однако настолько отличаются грамматически, что не являются взаимопонятными [З. М. Халилова, л. с.], что дает нам основание выделить инхокваринский в самостоятельный язык.

4.3.4.2. Даргинские языки

Так как даргинцы считают себя единым этносом, даргинские идиомы традиционно рассматриваются как диалекты единого «даргинского» языка. Однако, с чисто лингвистической точки зрения, многие даргинские «диалекты» невзаимопонятны друг с другом и могут считаться отдельными языками. В работе [Malyshev et al. 2019] сравнивается взаимопонимание между носителями кайтагского, кубачинского, цудахарского, муиринского и севернодаргинского (акушинское наречие) идиомов. Авторы приходят к выводу, что как минимум эти пять идиомов имеют настолько низкий уровень взаимопонимания, что должны считаться разными языками.

О том, что некоторые даргинские идиомы представляют собой самостоятельные языки, в разное время писали многие кавказоведы. Так, Е. А. Бокарев выделял кубачинский и кайтагский языки¹⁷ [Бокарев 1949]. С. М. Хайдаков выделял в отдельный язык речь села Мегеб [Хайдаков 1985]. М.-С. Мусаев считал кубачинский, кайтагский и чирагский близкородственными языками даргинской группы [Мусаев 1989]. А. А. Магометов выделял кубачинский язык [Магометов 1963]. В работах [Темирбулатова 2004] и [Гасанова 2011] обосновывается выделение кайтагского языка.

Революцию в даргиноведении совершил Ю. Б. Коряков, который начиная с 2000-ых гг. занимается классификацией даргинских идиомов на основе данных

¹⁷ Отметим, что в переписи 1926 года кубачинский и кайтагский языки числились отдельно.

лексикостатистики¹⁸ (см. Раздел 4.2.1). В разное время им выделялось 11 [Коруяков 2002], 17 [Коряков 2006; Коряков, Сумбатова 2007], 12-13 [Коряков 2012; Коруяков 2013] даргинских языков. По последним данным [Коряков 2021] выделяется 15 даргинских языков: севернодаргинский, цудахарский, усиша-бутринский, гапшиминский, муиринский, мегебский, юго-западно-даргинский, санжи-ицаринский, санакари-чахрижский, амузги-ширинский, кубачинский, аштынский, чирагский, кайтагский и шаринский. Однако по многим даргинским идиомам еще не собраны списки базовой лексики, поэтому данная классификация может измениться со временем.

Еще один подход к классификации даргинских идиомов был представлен осенью 2021 года в докладе Р. О. Муталова [Муталов 2021]. Вместе с лексикостатистическим методом Расул Османович предложил применять грамматический критерий разграничения идиомов, основанный на исследовании особенностей фонетики и морфологии. Предполагается, что, совместив данные двух подходов, можно будет определить более четкие границы между языком и диалектом. Для этого нужно было изучить, сколько в том или ином идиоме уникальных языковых явлений, которых нет в других (соседних) идиомах. Если таких явлений много, то идиом признавался языком (с оглядкой на лексикостатистику), если же число таких явлений незначительно, то идиом признавался диалектом (независимо от данных лексикостатистики). В результате анализа Расул Османович выделил 6 даргинских языков: акушинский, сирхя-цудахарский, кубачинский, мегебский, чирагский, кайтагский. Однако к данному подходу возник ряд методологических вопросов. Не совсем понятно, как отбирались признаки, был ли составлен единый набор признаков, релевантный для всех даргинских идиомов, что считалось одним признаком, а что — несколькими. Кроме того, довольно сложно формально определить, какого количества уникальных явлений достаточно, чтобы выделить идиом в отдельный язык. Самим Расулом Османовичем отмечается, что исследование еще не завершено, поэтому набор языков может измениться.

Несмотря на замечания к классификации Р. О. Муталова, мы считаем, что выделенные им языки вызовут меньше споров в научном и даргиноязычном сообществе, чем классификация Ю. Б. Корякова, поэтому акушинский, сирхя-цудахарский, кубачинский, мегебский, чирагский и кайтагский языки вошли в наш предварительный

¹⁸ Пробные лексикостатистические подсчеты в [Алексеев 1999: 156] показали, что фактически даргинский представляет собой целую группу языков, по глубине (минимальный ПС = 68%) сравнимую с германской, романской или славянской.

список. Отметим, что сейчас рассматриваются заявки на присвоение языковых кодов кубачинскому и кайтагскому в системе ISO 639-3¹⁹.

4.3.4.3. Лезгинские языки

Традиционно выделяется девять лезгинских языков, два из которых, кризский и будухский, не имеют значительного числа носителей на территории России, а представлены главным образом в Азербайджане. Еще четыре языка являются непроблемными с точки зрения вопроса “язык или диалект”: лезгинский, цахурский, арчинский и удинский языки. Кошанский “диалект” агульского языка имеет процент совпадений 88% с остальными диалектами, взаимопонимания между ними нет [Коряков 2020], поэтому мы выделяем кошанский в самостоятельный язык. В диссертации [Исрафилов 2007] приводятся некоторые сведения об особенностях фитинского говора агульского языка, однако на данный момент недостаточно данных для выделения его в самостоятельный язык. В цахурском языке дивергентным является гельмецкий диалект. В диссертации [Ибрагимов 2011] показано, что он является переходным звеном между цахурским и рутульским языками. Однако у нас нет данных лексикостатистики или сведений о взаимопонятности, поэтому пока мы не выделяем гельмецкий в отдельный язык. Многие исследователи, в частности, Я. Г. Тестелец, Т. А. Майсак, Ю. Б. Коряков, в устных беседах отмечали, что склоняются к разделению табасаранского языка на севернотабасаранский и южнотабасаранский из-за их невзаимопонятности, мы поддерживаем эту точку зрения. В рутульском языке выделяется дивергентный борчинско-хновский диалект, который не является взаимопонятным с мухадским диалектом, лежащим в основе литературного языка. В [Koryakov, Verhees 2021] говорится о потенциальной возможности разделить рутульский на собственно рутульский и южнорутульский языки, при этом под южнорутульским как раз подразумевается борчинско-хновский. Согласно лексикостатистике, процент совпадений между борчинско-хновским и другими рутульскими идиомами составляет 85-87% [Ю.Б. Коряков, л. с.]. Таким образом, борчинско-хновский мы также будем считать самостоятельным языком.

¹⁹ Заявка для кайтагского доступна по ссылке <https://iso639-3.sil.org/request/2021-008>. Заявка для кубачинского доступна по ссылке <https://iso639-3.sil.org/request/2021-024>.

4.3.4.4. Нахские языки

Среди нахских языков непроблемным оказывается ингушский язык. В чеченском языке проблему с точки зрения “язык или диалект” может представлять аккинско-орстхойское наречие, так как оно является переходным между чеченским и ингушскими языками [Коряков 2006: 26]. Так как чеченский и ингушский языки в принципе довольно близки друг другу, то лингвистические особенности указанного наречия не создают существенных проблем со взаимопонятностью. Поэтому аккинско-орстхойское наречие не выделяется нами в качестве самостоятельного языка из чеченского. Кроме того, в чеченском дивергентным является кистинский диалект, который по большей части распространен на территории Грузии. По нашим сведениям, компактных поселений кистинцев на территории России нет²⁰, поэтому мы не включаем кистинский в предварительный список.

4.3.5. Нивхский язык

В нивхском языке принято выделять три диалекта — амурский, восточносахалинский и северносахалинский²¹; значительное расхождение первых двух диалектов между собой позволяло исследователям рассуждать об их выделении в самостоятельные языки, см., в частности, [Панфилов 1968]. Что касается письменности, то в 1979 году было создано два алфавита — один для амурского и западносахалинского диалектов, другой для восточносахалинского диалекта [Gruzdeva, Janhunen 2018]. В сопоставительном словаре [Fortescue 2016] нивхский разделяется на три языка — амурский нивхский, восточносахалинский нивхский и южносахалинский нивхский; два нивхских языка (амурский и восточносахалинский) выделены также в [Comrie 1981: 266]. В [Gruzdeva 2021] было показано, что процент лексических соответствий между базовой лексикой амурского и восточносахалинского наречий составляет 84%. Северносахалинский вариант лингвистически ближе к амурскому. Таким образом, исходя из приведенных данных, в предварительный список языков России входят амурский нивхский и сахалинский нивхский.

²⁰ Отметим, однако, что по официальным результатам ВПН-2010 348 человек указали кистинский родным языком.

²¹ Ранее также выделялся южносахалинский диалект, носители которого после 1945 г. были переселены в Японию. Ныне южносахалинский диалект считается исчезнувшим [Gruzdeva, Janhunen 2018].

4.3.6. Тунгусо-маньчжурские языки

Непроблемными среди тунгусо-маньчжурских языков России являются негидальский, орокский (ульта), удэгейский, ульчский, эвенский и эвенкийский языки.

В нанайском языке проблемным является кур-урмийское наречие, которое совмещает в себе черты нанайского, эвенкийского, негидальского и удэгейского языков, поэтому некоторые исследователи ставили вопрос о том, считать ли его диалектом / наречием нанайского языка или отдельным языком, см. [Суник 1958: 9; Doerfer 1975]. В [Janhunen 2005: 42] кур-урмийский считается смешанным языком. В [Калинина, Оскольская 2016] сказано, что в последнее время (по-видимому, в отечественной лингвистике) кур-урмийский принято считать одним из диалектов нанайского языка. В настоящее время носителей кур-урмийского наречия, скорее всего, практически не осталось [В. С. Харитонов, л. с.; Оскольская 2017: 9], в силу чего нами было принято решение не выделять кур-урмийский нанайский в качестве отдельного языка.

4.3.7. Тюркские языки

Лингвистически многие тюркские языки крайне близки друг другу, что отражается также на лексикостатистических подсчетах — к примеру, процент лексических совпадений между татарским и башкирским языками составляет 95% [Коряков 2020]. Поэтому в рамках данной языковой семьи при принятии решения по проблемным идиомам мы во многом ориентировались на этническое самосознание носителей.

В ногайском языке обнаруживаются следующие проблемные идиомы: алабугатский, карагашский и юртовский.

Носители алабугатского идиома называют себя утарцами или ногайцами, однако в литературе принято обозначение этой этнической группы как алабугатских татар [Арсланов 2001]; во всероссийских переписях алабугатские татары включаются в группу астраханских татар. Язык алабугатских татар сформировался в результате смешения ногайцев с узбеками, туркменами, татарами и калмыками, поэтому некоторые исследователи считают алабугатско-татарский самостоятельным языком. В предварительном списке языков России мы выделяем этот язык как самостоятельный, указав в качестве основного названия “алабугатско-ногайский”.

Этническая группа карагашей в XVIII в. отделилась от ногайцев на Северном Кавказе и откочевала в астраханские степи, где оказалась в итоге в окружении калмыков, татар, русских и позднее — казахов [Сызранов 2008: 30]. В связи с этим у карагашей

возникла довольно сложная этническая и языковая идентичность, см. [Алексеев 2018]. Нечеткая этническая самоидентификация карагашей влечет за собой не вполне стабильную социолингвистическую ситуацию [Сатанова 2015]. Карагашский подвергся значительному влиянию со стороны казахского языка, в связи с чем некоторые исследователи выделяют карагашский как самостоятельный язык, см., в частности, [Арсланов 1991]. Мы также внесли карагашский в предварительный список языков России.

Что касается юртовцев, то еще в течение первой четверти XX века они имели ногайскую идентичность, а татарами стали считать себя после сталинской политики по укрупнению малых этнических групп; в результате сейчас одна часть юртовцев причисляет себя к татарам, а другая — к ногайцам [Ишмухамбетов 2020]. Юртовцы живут изолированно от остальных групп ногайцев (дагестанских и черкесских), на их язык существенно повлияли контакты с татарским языком [Арсланов 2005]. Нами было принято решение включить отдельный юртовско-ногайский язык в предварительный список языков России.

Крымчакский в силу значительного влияния со стороны крымскотатарского языка иногда рассматривается как его этнолект (так сделано, к примеру, в [Polinsky 1992]). Однако крымчаки осознают себя как отдельную этническую группу. Кроме того, в системе стандартизации ISO 639-3 крымчакский имеет отдельный языковой код. Судя по полученным нами данным, носителей крымчакского языка не осталось в России (см. подробнее раздел 6), поэтому крымчакский входит в список заснувших языков России (см. Таблицу 4 в Приложении).

Диалекты караимского языка — тракайский (диалект литовских караимов), галичский (диалект караимов Западной Украины) и крымский — иногда рассматриваются как три самостоятельных языка. Так делается, в частности, в “Атласе языков мира, находящихся под угрозой исчезновения”, составленном ЮНЕСКО. Тем не менее, в системе стандартизации ISO 639-3 для караимского выделяется только один языковой код. Караимский идиом, распространенный в России в Крыму, подвергся контактам с крымскотатарским языком и сблизился с ним; в частности, в [Polinsky 1992] на основании лексикостатистических подсчетов делается вывод, что крымский караимский представляет собой этнолект крымскотатарского языка (об истории классификации караимского идиома, бытующего в Крыму, как языка или этнолекта см. в [Ельяшевич 2020]). Так как караимы выделяют себя как отдельную этническую группу

и караимскому языку присвоен языковой код в ISO 639-3, мы склоняемся к тому, чтобы считать караимский самостоятельным языком.

Для сибирскотатарского языка выделяется отдельный индекс в ISO 639-3; кроме того, сибирские татары осознают себя как отдельную группу татар. Что касается самостоятельной литературной нормы, то можно сказать, что она пока не сформирована, хотя в этом направлении ведется работа филологом М. А. Сагидуллиным, см., к примеру, [Сагидуллин 2008], а также изданный им русско-сибирскотатарский словарь [Сагидуллин 2010]. В школе детям-носителям сибирскотатарского преподают литературный татарский язык, что в силу отличия их идиома вызывает существенные проблемы, см., в частности, [Хомякова 2021]. Таким образом, мы выделяем сибирскотатарский как самостоятельный язык.

Внутри сибирскотатарского языка проблемным является идиом барабинских татар. В некоторых исследованиях он выделяется как отдельный язык (см., к примеру, [Селютина и др. 2013]). В 2018 году Совет при Президенте Российской Федерации по развитию гражданского общества и правам человека по итогам 26-го выездного заседания в Новосибирской области выпустил рекомендацию, в которой, среди прочего, Губернатору и Правительству Новосибирской области рекомендуется обратиться в Правительство РФ с инициативой признания барабинских татар коренным малочисленным народом Сибири и Дальнего Востока, а также поручить Министерству образования Новосибирской области разработать азбуки и учебники для преподавания барабинского языка в детских садах и школах [Рекомендации СПЧ 2018]²². Барабинский включен как самостоятельный язык в “Атлас языков мира, находящихся под угрозой исчезновения” ЮНЕСКО. Таким образом, мы также приняли решение выделять барабинский татарский как отдельный язык.

Еще одним проблемным сибирскотатарским идиомом является калмакский. Некоторые исследователи показывают, что калмакский по лингвистическим данным можно считать самостоятельным языком, см. [Уртегешев 2016]. Сами калмаки указывают на значительное отличие их языка от языка казанских татар, а также соседних тюркоязычных групп (эуштинцев и чатов); при этом мнение о том, является ли их язык самостоятельным, среди калмаков разнится [Уртегешев 2016]. Язык калмаков можно причислить к исчезающим — носителями являются примерно 20 человек [Н. С.

²² В этом документе фигурирует также чатский язык, однако пока у нас нет достаточно оснований рассматривать данный идиом как самостоятельный язык.

Уртегешев, л. с.]. Ввиду малочисленности представителей данного идиома, а также того, что существуют исследователи, считающие данный идиом самостоятельным языком, мы приняли решение включить калмакский язык в предварительный список языков России.

Перейдем к обсуждению идиомов, которые традиционно считаются диалектами единого алтайского языка. Данные лексикостатистики показывают, что кумандинский, тубаларский и челканский разошлись настолько, что их следует считать самостоятельными языками [Коряков 2020]. Кроме того, носители данных идиомов осознают себя как отдельные этнические группы, поэтому мы разделяем североалтайские идиомы на три самостоятельных языка — кумандинский, тубаларский и челканский. Что касается южноалтайских идиомов, то лексикостатистически существенно отличается телеутский язык (88% совпадений с остальными южноалтайскими идиомами), тогда как процент совпадений в базовой лексике между алтай-кижи и теленгитским составляет 93% [Ю. Б. Коряков, л. с.]. Однако теленгиты считают себя отдельной этнической группой. Таким образом, мы выделяем телеутский, теленгитский и алтай-кижи как отдельные языки.

4.3.8. Уральские языки

В языках уральской семьи, распространенных на территории России, идиомы, проблемные с точки зрения вопроса “язык или диалект”, можно найти во всех языковых группах.

4.3.8.1. Марийские языки

Лугово-восточный марийский (луговомарийский) и горномарийский идиомы в научной литературе считаются либо наречиями марийского языка, либо самостоятельными языками. Мы склонны считать их самостоятельными языками. Хотя марийские идиомы в целом близки друг к другу, они обладают целым рядом отличий на всех уровнях языка, что приводит к нарушению взаимопонятности (см., к примеру, [Коростелев 2003; Кашкин, Мордашова 2021]). Дополнительным фактором для придания лугово-восточному марийскому и горномарийскому идиомам статуса языка является наличие у них отличных друг от друга литературных норм. Кроме того, как показано в [Шабыков, Кудрявцева, Зорина 2020], горные мари обладают высоким уровнем этнического самосознания. Таким образом, лугово-восточный марийский и

горномарийский языки оказываются включенными в наш предварительный список языков России.

Статус северо-западного наречия марийского языка пока остается спорным. Лингвистически оно обладает чертами как луговомарийских, так и горномарийских идиомов, но ближе оказывается к последним. В [Юзыкайн 2000] содержатся обрывочные сведения о том, что у носителей северо-западного наречия имеются трудности с пониманием других марийских диалектов; там же сообщается, что в школах им преподают лугово-восточный марийский литературный язык, а не горномарийский, к которому диалекты северо-западного наречия лингвистически более близки. В 1995 году был издан букварь на северо-западном наречии [Дмитриев, Дмитриева, Тужаров 1995]. В основном носители северо-западного наречия проживают вне Республики Марий Эл — в Кировской и Нижегородской областях, где марийские языки не являются государственными. Для улучшения языковой ситуации данного идиома можно было поднять его статус с наречия до языка, однако для этого необходимо провести детальные социолингвистические исследования, особенно касающиеся степени взаимопонятности. Дело в том, что в 2019 году в организацию по стандартизации ISO 639-3 была подана заявка о присвоении северо-западному наречию марийского отдельного языкового кода, которая, однако, была отклонена в том числе на основании недостатка данных о нарушенной взаимопонятности с остальными марийскими идиомами²³.

4.3.8.2. Мордовские языки

В мокшанском языке нами не было зафиксировано каких-либо диалектов, значительно отошедших от других. В эрзянском языке выделяется один изолированный диалект — шокшинский, распространенный на северо-западе Республики Мордовия в Теньгушевской районе. Отмечается его переходность, так как этот диалект исторически оказался в изоляции от других эрзянских говоров и в окружении мокшанских диалектов [Феоктистов 2001]. Тем не менее, нами не было найдено источников, которые бы отмечали нарушенную взаимопонятность между носителями шокшинского диалекта и других эрзянских диалектов, поэтому мы не выделяем шокшинский как самостоятельный язык.

²³ С текстом заявки и причинами ее отклонения можно ознакомиться на следующей странице: <https://iso639-3.sil.org/request/2019-001>.

4.3.8.3. Пермские языки

Среди удмуртских идиомов проблемным с точки зрения “язык или диалект” представляется нам бесермянское наречие, так как бесермяне обладают отчетливым этническим самосознанием, противопоставляя себя остальным удмуртам; кроме того, как указывается в [Шкляев 1997], большая половина бесермян считают свой идиом самостоятельным языком. Бесермяне входят в Единый перечень коренных малочисленных народов Российской Федерации. Тем не менее, существенных отличий бесермянского наречия от прочих удмуртских идиомов не обнаруживается²⁴. Таким образом, мы не выделяем бесермянский в качестве отдельного языка.

В коми-зырянском языке выделяется коми-ижемский диалект, носители которого также выделяют себя в особую этническую группу. Сведений о нарушенной взаимопонятности между носителями коми-ижемского и носителями остальных диалектов коми нам обнаружить не удалось, поэтому мы не склонны считать коми-ижемский отдельным языком.

Отдельной от коми-пермяков этнической группой себя считают коми-язьвинцы. В работе [Норманская 2020] было показано, что на основании фонетических и этимологических изменений коми-язьвинский можно считать отдельным языком. Кроме того, на коми-язьвинском в 2000-х гг. был выпущен букварь и книга для чтения [Паршакова 2003, 2008]. Поэтому мы будем считать коми-язьвинский отдельным языком наряду с коми-пермяцким.

Еще один проблемный идиом, который можно выделить среди коми-зырянских диалектов, — это удорский диалект, подвергшийся некоторому прибалтийско-финскому влиянию (об этом см. подробнее [Федюнева 2008]). Однако, этот факт, по-видимому, существенно не влияет на степень взаимопонятности с остальными коми-зырянскими идиомами, поэтому удорский не был включен в предварительный список языков России.

4.3.8.4. Прибалтийско-финские языки

Среди прибалтийско-финских языков проблемными являются ижорский, вепсский и водский языки.

В карельском языке выделяют три наречия — собственно карельское, ливвиковское и людиковское, между которыми отмечается нарушение

²⁴ См. также мнение одного из носителей бесермянского наречия по отношению к взаимопонятности с другими удмуртскими диалектами: <https://realnoevremya.ru/articles/177324-my-besermyane-pervymi-poteryaem-svoy-yazyk>.

взаимопонятности (см. [Кондрашкина 2016; Влахов и др. 2020]). В системе стандартизации ISO 639-3 имеются отдельные языковые коды для карельского, ливвиковского и людиковского идиомов. Кроме того, сами карелы разделяют себя на основании различий в своей речи [Рягоев 1993]. Исходя из перечисленных фактов мы также выделяем собственно карельский, ливвиковский карельский и людиковский карельский языки. Среди собственно карельских диалектов лингвистически выделяется тверской диалект. Носители тверского карельского проживают отдельно от носителей остальных карельских идиомов. В 1996 году был создан отдельный алфавит для тверского карельского, однако в 2007 году Правительством Республики Карелия был утвержден алфавит карельского языка, единый для всех диалектов [Постановление Правительства Республики Карелия 2007]. Мы считаем важным поддержать развитие варианта карельского языка, распространенного в Тверской области в силу географической обособленности его носителей²⁵, поэтому выделяем его в нашем предварительном списке как самостоятельный язык. Отметим, что в 2016 году Совет при Президенте Российской Федерации по развитию гражданского общества и правам человека опубликовал текст рекомендаций, в которых в том числе уделяется внимание положению тверского карельского в области [Рекомендации СПЧ 2016]. Рекомендации касаются придания тверскому карельскому языку официального статуса, подготовки учителей тверского карельского и развитию литературы на тверском карельском.

Среди эстонских идиомов выделяется сето (или сету), который чаще всего считается диалектом или говором вырусского диалекта²⁶ эстонского языка. Сето считают себя отдельной этнической группой, и в 2010 году народ сето был внесен в Единый перечень коренных малочисленных народов Российской Федерации. Сведения о значительных расхождениях языка сето и литературного эстонского содержатся в [Chalvin 2011: 128-129; Агранат 2021]; в [Pajusalu 2003: 232] проводится сравнение с водским языком, который имеет по сравнению с эстонским примерно столько же отличий, что и сето, но при этом считается отдельным языком. В 2011 году был выпущен букварь на сето, и язык вскоре стал преподаваться в компактных поселениях сето в Эстонии [Агранат 2019]. Сето, в отличие от большинства эстонцев, исповедуют православие. На основании указанной информации мы также принимаем решение о том,

²⁵ Фактор географической обособленности в случае карельских идиомов нам представляется крайне важным. До недавнего времени обособленно от карел в Карелии проживали так называемые валдайские карелы в Новгородской области, однако в 1980-1990-х гг. их идиом был утрачен [Пунжина (сост.) 2001: 6].

²⁶ В настоящее время в Эстонии вырусский признан отдельным языком.

что сето может считаться отдельным языком. Сето также проживают в Красноярском крае, куда они переселились на рубеже XIX-XX во время аграрной реформы П. А. Столыпина. Язык сибирских сето отличается некоторой вариативностью, так как сето переселялись из разных мест и, соответственно, являлись представителями различных говоров [Агранат 2020]. Сибирские сето сохраняют связь с исторической родиной; этот факт позволяет нам предположить, что взаимопонятность между идиомами сето, проживающими в Сибири и в Эстонии, сохраняется. Поэтому оснований выделять язык сибирских сето как самостоятельный, помимо их обособленного проживания, нет.

4.3.8.5. Саамские языки

В зарубежной лингвистике саамские идиомы уже достаточно давно считаются самостоятельными языками, в системе стандартизации ISO 639-3 им также присвоены отдельные коды. В отечественных работах, напротив, саамские идиомы зачастую рассматриваются как единый язык с множеством разошедшихся диалектов. В [Sammalahti 1998] сообщается, что лексическое сходство саамских идиомов составляет от 80% до свыше 90% (напомним, что, согласно подсчетам Ю. Б. Корякова [2017], условной границей между языком и диалектом можно считать порог в 91%). На территории России представлены четыре саамских идиома — кильдинский, йоканьгский, скольтский и бабинский²⁷. В [Керт 1966] сообщается, что представители трех первых перечисленных нами диалектов долгое время не имели контакта друг с другом. В базе данных ASJP имеются списки из 40 наиболее устойчивых слов списка Сводеша для бабинского саамского, кильдинского саамского и скольтского саамского. Для этих трех идиомов нами было подсчитано расстояние LDN²⁸ (см. раздел 4.2.1); результаты подсчета приведены в Таблице 1²⁹. Расстояние в 0,51, согласно [Wichmann 2019], является пороговым значением для различения языка и диалекта — если расстояние между некоторыми двумя идиомами превышает это пороговое значение, то это говорит о том, что данные идиомы разошлись настолько, что их можно считать отдельными языками. Как видно из приведенной таблицы, между всеми парами идиомов расстояние LDN составляет больше 0,51. Таким образом, мы предполагаем, что

²⁷ Бабинский саамский мы считаем заснувшим языком, см. подробнее раздел 6.

²⁸ По сравнению с метрикой LDND, для нее определено точное пороговое значение, касающееся различения языка и диалекта.

²⁹ В проекте ASJP были разработаны специальные программы для подсчета различных метрик, см. <https://asjp.cild.org/software>. Мы использовали программу для подсчета матриц расстояния, написанную на языке программирования R.

кильдинский саамский, скольтский саамский, йоканьгский саамский и бабинский саамский можно считать самостоятельными языками.

Таблица 1. Расстояние LDN между парами саамских идиомов (кильдинский саамский = KILDIN_SAAMI, скольтский саамский = SKOLT_SAAMI, бабинский саамский = AKKALA_SAAMI); списки базовой лексики взяты из базы данных ASJP.

	AKKALA_SAAMI	KILDIN_SAAMI	SKOLT_SAAMI
AKKALA_SAAMI	0.0000		
KILDIN_SAAMI	0.5694	0.0000	
SKOLT_SAAMI	0.5432	0.5324	0.0000

4.3.8.6. Самодийские языки

Автоматически в предварительный список языков России был внесен нганасанский язык как непроблемный.

В ненецком языке традиционно выделяют два наречия — тундровое и лесное, между которыми даже в энциклопедических источниках начиная со статьи Н. М. Терещенко в “Языках народов СССР” [1966] отмечается нарушение взаимопонятности. Лексикостатистические подсчеты для ненецких идиомов, представленные в [Коряков 2018a], показывают, что процент лексических совпадений между тундровыми и лесными диалектами составляет в среднем 88,4%. Согласно получившемуся значению, тундровый и лесной ненецкий могут считаться отдельными языками. Однако отметим, что в системе стандартизации ISO 639-3 они все еще не получили отдельных языковых кодов. На основании приведенных фактов мы разделяем ненецкий на тундровый ненецкий и лесной ненецкий.

Для лесного энецкого и тундрового энецкого, напротив, имеются отдельные коды в ISO 639-3, хотя в литературе содержатся противоречивые данные о их взаимопонятности (ср., к примеру, [Урманчиева 2002; Ханина, Шлуинский 2016], одной стороны, и [Salminen 1997; Siegl, Rießler 2015], с другой стороны). Кроме того, как показано в [Коряков 2018a], процент лексических совпадений в базовой лексике между лесным и тундровым энецким составляет 93,9% (или 92,9%, если не считать местоимение 2 л. ед.ч. заимствованным), что выше границы между языком и диалектом.

Что касается этнической самоидентификации, то лесные и тундровые энцы не имеют общего самоназвания: лесные энцы именуют себя бай, а тундровые — маду (или мадду). Таким образом, учитывая различное этническое самосознание и наличие данных о нарушенной взаимопонятности, мы будем считать лесной энецкий и тундровый энецкий самостоятельными языками.

Что касается селькупского языка, то подавляющее большинство отечественных исследователей говорят о едином селькупском языке. В [Janhunen, Salminen 1993] постулируется существование трех селькупских языков — северноселькупского, центральноселькупского и южноселькупского. С 1980-х гг. пытаются возродить селькупскую письменность, которая была основана на северном наречии; как отмечается в [Казакевич 2016], для южных селькупов учебники на этой письменности были совершенно непригодны в силу разительных отличий между диалектами; о неудобстве использования литературы, базирующейся на северном наречии, упоминается также в справочнике Ethnologue³⁰. Самоназвания различных групп селькупов не совпадают, см. [Хелимский 2002а]. По лексикостатистическим подсчетам, сделанным в [Коряков 2018а], северноселькупский может считаться самостоятельным языком. Таким образом, мы также разделяем селькупский на северноселькупский и южноселькупский языки³¹.

4.3.8.7. Угорские языки

Диалектная раздробленность хантыйского языка представляется крайне высокой. Распространены два подхода к выделению диалектных массивов в хантыйском: один подход выделяет три наречия — восточное, южное и северное [Хонти 1993; Хонти, Николаева, Каксин 2005]; согласно второму подходу, в хантыйском имеются западное и восточное наречия [Хелимский 2002б]. При втором подходе южное и северное наречия объединяются в западное. В исследовании хантыйских диалектов [Соловар, Нахрачева, Шиянова 2016] говорится, что “различия [между западным и восточным наречиями] настолько велики, что если границы между языком и диалектом определять только лингвистически на уровне языкового сознания говорящих и взаимного понимания- непонимания, то следует говорить о двух языках — западных и восточных ханты”. В

³⁰ См. <https://www.ethnologue.com/language/sel/24>.

³¹ Название последнего представляет некоторую терминологическую сложность. В списке языков [Коряков и др. 2020] принято в качестве основного названия “язык селькупов Томской области”. Среди прочих вариантов названий в [Коряков 2018а] перечисляются следующие: центрально-южноселькупский, язык южных селькупов, томско-селькупский, селькупско-остяцкий.

[Кошкарева 2016] отмечается недостаточная изученность диалектной ситуации хантыйского языка. Так как для восточного и западного наречия в литературе четко говорится о нарушенной взаимопонятности, то мы выделяем восточнохантыйский (или кантыкский) в качестве самостоятельного языка. Отметим также, что в [Коряков 2017] поднимается вопрос о единстве восточнохантыйского (кантыкского) языка: сообщается, что между крайними членами диалектного континуума процент лексических совпадений составляет 86-88%. Южнохантыйские идиомы к настоящему времени являются, по-видимому, исчезнувшими [А. Ю. Фильченко, л. с.]. Согласно подсчетам по базе данных Starling³², проценты лексических совпадений между южнохантыйскими (верхнедемьянским, кондинским и низямским) и другими хантыйскими диалектами составляют менее 91%³³, поэтому мы разделяем севернохантыйский и южнохантыйский на два самостоятельных языка.

Диалекты мансийского языка подразделяются на северные, южные, восточные и западные, из которых к настоящему времени сохранились только первые. В [Коряков 2017] на основании лексикостатистических подсчетов, сделанных по спискам базовой лексики из словаря [Munkácsi, Kálmán 1986] делается вывод, что можно выделить два языка — северномансийский и исчезнувший южно-средне-мансийский. Однако в недавнем исследовании [Федотова 2021], учитывающем лексические данные, полученные при полевой работе в XXI веке, также на основе совпадения в базовой лексике постулируется единство мансийского языка. При обоих подходах признается существование в наше время только одного мансийского, поэтому именно он и вошел в наш предварительный список. Учитывая более новые данные, изложенные в [Федотова 2021], мы не включаем в список заснувших языков мансийские идиомы (как это сделано в списке проекта “Языки России” [Коряков и др. 2020]).

4.3.9. Чукотско-камчатские языки

Ительменский, корякский и чукотский языки не вызывают вопросов с точки зрения составления списка языков России. Проблемным является алюторский язык.

³² В базе данных Starling имеются 100-словные списки для обско-угорских языков, см. <https://starlingdb.org/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100\ura\oug&limit=-1>.

³³ Отметим, однако, что процент лексических совпадений между низямским и остальными южнохантыйскими идиомами — кондинским и верхнедемьянским — составляет 89% и 88% соответственно. Для низямского диалекта отмечается его переходность по отношению к северному наречию [Хелимский 20026].

Алюторский — язык оседлых коряков-нымыланов — раньше считался диалектом корякского языка [Жукова 1997], однако Ю. Нагаяма отмечает, что носители корякского и алюторского языков понимают друг друга с большим трудом [Ю. Нагаяма, л. с.]. Термин «алюторский» происходит от названия бывшего села Алют. К алюторскому в широком смысле, помимо собственно алюторского, относят и близкие к нему паланский, карагинский, рекинниковский диалекты. Все группы коряков, владеющие данным идиомом, предпочитают называть его нимыланским [Мальцева 2016]. В силу ряда причин алюторский не был представлен в качестве отдельного языка в энциклопедии “Языки Российской Федерации и соседних государств” [2001], хотя был описан в энциклопедии “Языки народов СССР” [1968], а также в недавней энциклопедии “Язык и общество” [2016]. Мы включаем алюторский язык в наш предварительный список.

4.3.10. Эскимосско-алеутские языки

Среди эскимосско-алеутских языков проблемными являются науканский и чаплинский (уназикский) языки. В 1960-70-е гг. они считались диалектами эскимосского языка, однако более тщательные исследования выявили существенные различия между ними [Вахтин 2016], позволяющие выделить их в самостоятельные языки, что подтверждается также данными лексикостатистики (82% по [Коряков 2020]).

В 1930-е гг. в основу литературной нормы был положен чаплинский диалект. Науканский литературной нормы не имеет. По экспертным оценкам 2016 года, носителей науканского осталось меньше 100 человек. Оба языка относятся к исчезающим [Вахтин 2016]. Кроме того, каждый из представленных языков имеет отдельный код в системе стандартизации ISO 639-3. При этом, согласно шкале витальности ЮНЕСКО, науканский язык находится в критическом состоянии, а чаплинский — под серьезной угрозой. Учитывая всё вышеперечисленное, мы включаем науканский и чаплинский языки в предварительный список.

4.3.11. Юкагирский язык

Юкагирский язык представлен в настоящее время двумя идиомами — северным (тундренным) и южным (лесным). Значительные расхождения между этими двумя идиомами отмечались многими исследователями, см. [Николаева, Хелимский 1996; Курилов 2003: 9; Maslova 2003: 1]. Данные по лексикостатистике несколько разнятся. По подсчетам М. А. Живлова, процент лексических совпадений по 100-словному списку между северным юкагирским и южным юкагирским составляет 47% [Немировский

2017]. В [Коряков 2020] сообщается, что уровень лексических совпадений в базовой лексике оценивается в 74%. Так или иначе, все эти значения существенно ниже порога в 91% между языком и диалектом. Письменность существует и для северного, и для южного юкагирского. Поэтому мы выделяем северноюкагирский и южноюкагирский как два самостоятельных языка.

5. Экстратерриториальные языки России

Другой важной проблемой является вопрос о том, какие языки, носители которых присутствуют на территории РФ, следует считать языками России. В последней переписи населения, прошедшей в 2010 году, в итоговой таблице “Владение языками населением Российской Федерации” фиксируется 172 языка. В их число входят, например, такие “экзотические” языки, как амхарский (311 владеющих), бенгали (393 носителей), индонезийский (515 владеющих), исландский (233 владеющих) и некоторые другие. Радикальным решением данной проблемы могло бы явиться включение всех языков, зафиксированных переписью, в список языков России. Такое решение представляется нам не очень осмысленным в силу того, что в переписи имеют право участвовать граждане, приехавшие в РФ на обучение или на работу сроком более одного года, — группу таких людей сложно назвать “народом России”. Поэтому разумнее использовать критерии, отражающие традиционность и длительность проживания некоторой этнической группы на территории России. Очевидно, что в списке языков России должны быть зафиксированы языки народов, включенных в Единый перечень коренных малочисленных народов Российской Федерации [Постановление 2000а]. Кроме того, в список языков России логично включать языки, чей основной ареал затрагивает территорию России (например, удмурты не фигурируют в Едином перечне коренных малочисленных народов РФ в силу своей многочисленности, однако подавляющее большинство удмуртов проживают на территории России). Проблемными представляются языки, основная масса носителей которых проживает вне территории РФ. Так, в России довольно многочисленно финноязычное население, однако большая часть финнов проживает в Финляндии.

Языки, чей основной ареал находится за пределами РФ, мы для краткости будем называть экстратерриториальными. Проанализировав данные двух последних всероссийских переписей населения, мы составили список из более 150 экстратерриториальных языков. Однако в наш список языков России были включены

только те языки, которые отвечают ряду параметров. Среди возможных критериев, согласно которым экстратерриториальные языки могут быть включены в список языков России, можно назвать следующие факторы:

- 1) Упоминание языка в законодательстве РФ, наличие у него какого-либо статуса;
- 2) Наличие языка в системе образования РФ (в качестве языка преподавания, школьного предмета, факультатива, предмета в воскресной школе и т. п.);
- 3) Традиционность проживания носителей языка на нынешней территории РФ³⁴;
- 4) Наличие на территории России компактных поселений носителей языка³⁵;
- 5) Пересечение границы ареала распространения языка государственной границы РФ;
- 6) Развитие на территории России особого диалекта языка;
- 7) Наличие у диаспоры общественных объединений на территории РФ.

Кроме того, на проблему экстратерриториальных языков можно смотреть сквозь призму ревитализации и поддержки. Например, абхазский язык обычно не включается в список языков России (см. [Солнцев, Михальченко (отв. ред.) 2000, 2003; Михальченко (гл. ред.) 2016; Коряков и др. 2020]), однако по шкале витальности ЮНЕСКО он имеет статус уязвимого. Включение абхазского языка в список языков России могло бы привлечь к нему внимание общественности и лингвистов. Поэтому мы обращали внимание также на следующие факторы:

- 8) Статус языка по шкале витальности, разработанной ЮНЕСКО;
- 9) Наличие у языка какого-либо официального статуса в других странах;
- 10) Наличие у языка письменности и литературной нормы.

Последовательно рассмотрим, какие языки были включены в наш предварительный список и на каком основании.

В итоговый список прежде всего были включены языки, фигурирующие в российском законодательстве. К таким языкам можно отнести азербайджанский, казахский, украинский и финский языки. Азербайджанцы постановлением от 2000 г. включены в перечень коренных малочисленных народов Республики Дагестан

³⁴ Критерий традиционности проживания нуждается, однако, в установлении порогового значения — длительности проживания носителей языка на территории России.

³⁵ Для критерия компактности проживания также необходимо подбирать пороговое значение. Так, в законодательстве ряда западноевропейских государств язык некоторого национального меньшинства получает официальный статус в регионах страны, где носители этого языка составляют более 15% или 20% населения; см., к примеру, законодательство Словении, где разрешается использовать язык национального меньшинства на официальном уровне, если численность его носителей в пределах муниципального образования составляет 15% и более [Predpis 1999].

[Постановление 2000б]. В Законе Республики Алтай использование казахского языка в официальных средствах общения разрешается в местах компактного проживания его носителей [Закон Республики Алтай 1993, ст. 4]. Использование финского языка в официальных сферах регулируется Законом Республики Карелия [Закон Республики Карелия 2004]. Несмотря на то, что азербайджанский, казахский и финский языки упоминаются в законодательстве указанных регионов России, их статус четко нигде не прописывается. Для перечисленных выше языков статус указывается только для украинского языка, который является одним из государственных языков Республики Крым [Конституция Республики Крым 2014, ст. 10, п. 1].

Помимо проблемы нечетко прописанного статуса, можно также указать на то, что некоторые языки не упоминаются в законодательстве тех регионов, где представлена значительная часть их носителей. Так, казахский язык фигурирует в законодательстве Республики Алтай, однако отсутствует в законодательстве Астраханской области, в которой имеется наиболее крупная в России община казахов (для сравнения, по данным последней переписи 2010 г., казахи составляют 16,3% населения Астраханской области и около 6% — Республики Алтай). Такое положение связано с тем, что, согласно законодательству РФ, только республики могут устанавливать официальные языки [Закон о языках 1991, ст. 3, п. 2].

Решения по остальным экстратерриториальным языкам принимались по совокупности различных критериев, с учетом того, что каждый из них имеет разный так называемый “вес”. Кратко перечислим языки, которые вошли в наш предварительный список языков России, с указанием основных аргументов в пользу такого включения.

- Абхазский — 1. У абхазов есть небольшие компактные поселения на территории России; 2. Ареал распространения абхазского языка переходит границу РФ; 3. Абхазский язык имеет уязвимый статус по шкале витальности ЮНЕСКО.
- Белорусский — 1. У белорусов есть компактные поселения на территории России; 2. Ареал распространения белорусского языка переходит границу РФ; 3. Белорусский язык имеет уязвимый статус по шкале витальности ЮНЕСКО; 4. До недавнего времени белорусский язык присутствовал в системе школьного образования в качестве предмета.
- Корейский — 1. У корейцев есть компактные поселения на территории России; 2. Ареал распространения корейского языка переходит границу РФ; 3. Корейский язык присутствует в системе школьного образования в качестве предмета и факультатива.

- Нижненемецкий — 1. У немцев есть компактные поселения в Сибири, ранее в Поволжье и на Кавказе; 2. Нижненемецкий изучается в детских клубах на Алтае.
- Верхненемецкий — 1. У немцев есть компактные поселения в Сибири, ранее в Поволжье и на Кавказе; 2. Верхненемецкий язык присутствует в школьном образовании в качестве предмета.
- Армянский — 1. У армян есть компактные поселения на территории России; 2. Армяне проживают на территории России с XIX века; 3. Армянский язык присутствует в школьном образовании в качестве предмета и факультатива.
- Таджикский — 1. У таджиков есть компактные поселения на территории России; 2. Таджикский язык изучается на территории России как предмет в воскресных школах; 3. В России есть частные детские сады с обучением на таджикском языке.
- Цыганский — 1. У цыган есть компактные поселения в Ставропольском крае, Ростовской области; 2. В некоторых начальных школах цыганский язык присутствует в качестве языка обучения; 3. Цыганский язык имеет статус “есть угроза исчезновения” по шкале ЮНЕСКО.
- Польский — 1. У поляков есть компактные поселения в Сибири; 2. Польский язык присутствует в системе школьного образования в качестве предмета и факультатива.
- Чешский — 1. У чехов есть компактные поселения в Крыму; 2. Чехи проживают в России с 60-х гг. XIX века.
- Новогреческий — 1. У фракийских греков, которые являются носителями фракийского говора северного наречия новогреческого языка, имеется компактное поселение в с. Чернополье Республики Крым; 2. Язык присутствует в системе школьного образования в России; 3. Понтийские греки получают образование на новогреческом.
- Понтийский — 1. Понтийские греки переселились на территорию России в XIX-XX вв.; 2. У понтийских греков есть компактные поселения на Северном Кавказе.
- Курманджи — 1. У курдов есть компактные поселения на юге и в центре европейской части России.
- Литовский — 1. У литовцев есть компактные поселения в Калининградской области; 2. Ареал распространения литовского языка переходит границу РФ; 3. Литовский язык представлен в школьном образовании на территории России.
- Болгарский — 1. Болгары традиционно проживают в Крыму; 2. Указом президента РФ предусмотрено содействие получению гражданами Российской Федерации,

проживающими на территориях Республики Крым и г. Севастополя, основного общего образования в том числе на болгарском языке; 3. Присутствует в системе школьного образования.

- Латышский — 1. У латышей есть небольшие компактные поселения в Сибири; 2. Ареал распространения латышского языка переходит границу РФ; 3. Латышский язык представлен в школьном образовании России.
- Латгальский — 1. У латгальцев есть компактные поселения (в Новосибирской области, Красноярском крае и Кемеровской области); 2. Ареал распространения латгальского языка затрагивает границу РФ; 3. Латгальцы проживают в России с конца XIX века.
- Молдавский — 1. У молдаван есть компактные поселения (например, село Молдаванское в Краснодарском крае); 2. Молдавский язык представлен в системе школьного образования России в качестве предмета.
- Грузинский — 1. У грузин есть компактные поселения в Северной Осетии-Алании; 2. Грузинский язык представлен в школьном образовании России в качестве предмета и факультатива.
- Идиш — 1. Идиш представлен в языковом ландшафте Еврейской автономной области³⁶; 2. До недавнего времени идиш был представлен в школьном образовании России.
- Монгольский — 1. У монголов есть ареал постоянного проживания на юге Тувы; 2. Ареал распространения монгольского языка переходит границу РФ.
- Ассирийский новоарамейский — 1. У ассирийцев есть компактные поселения; 2. Ассирийский новоарамейский язык представлен в системе школьного образования России в качестве факультатива.
- Китайский — 1. Китайский является традиционным языком малых народностей тазов и гольдов в Приморье; 2. Тазовский вариант китайского языка находится на грани исчезновения.
- Дунганский — 1. Дунгане компактно проживают в селах Привольное, Кочетное и Скатное Ровенского района Саратовской области.
- Узбекский — 1. Узбеки живут компактно в нескольких деревнях в разных регионах России.

³⁶ Отметим, что идиш не имеет официального статуса в Еврейской автономной области (ст. 6 в [Устав ЕАО]).

- Киргизский — 1. У киргизов есть компактные поселения на территории России; 2. Киргизский язык присутствует в системе образования России: изучается как предмет в воскресных школах, есть частные детские сады с обучением на киргизском языке.
- Туркменский — 1. У туркмен есть компактные поселения в Астраханской области и в Ставропольском крае; 2. Туркменский язык присутствует в системе школьного образования России.
- Гагаузский — 1. В России проживает весьма значительная численность гагаузов; 2. В России имеются компактные поселения гагаузов; 3. Есть частная организация, занимающаяся преподаванием гагаузского языка.
- Турецкий — 1. Турки-месхетинцы начали переселяться на территорию современной России в 1980-1990-х гг.; 1. У турок есть компактные поселения в Ставропольском крае, Ростовской области, Краснодарском крае, Калмыкии, Кабардино-Балкарии.
- Татский — 1. Предположительно есть компактные поселения (см. раздел 4.4.3); 2. Имеет статус “серьезная угроза исчезновения”; 3. Язык не является официальным ни в одном государстве.
- Удинский — 1. Значительные группы удин проживают в Ростовской области, в Краснодарском крае, Волгоградской области; 2. Язык не является официальным ни в одном государстве.

6. Заснувшие языки

Проблему для составления списка языков России представляют языки, исчезнувшие на территории России, а также заснувшие языки. Под заснувшими языками мы подразумеваем не только языки, носителей которых не осталось, но и такие языки, у которых остались только пассивные носители, не использующие язык в повседневном общении, хотя обладающие некоторым уровнем знания языка. Проблемы для составления списка языков России возникают по нескольким причинам. Во-первых, имеющиеся социолингвистические данные не всегда дают представление о том, действительно ли в России не осталось больше носителей предположительно исчезнувшего идиома. Во-вторых, важно учитывать нацеленность сообщества на восстановление языка. Мероприятия по ревитализации языка могут перевести язык из категории заснувших в категорию функционирующих языков. На возможность ревитализации языка влияет степень его задокументированности.

Нами был составлен список языков, исчезнувших со второй половины XX века. Эти языки можно разделить на несколько групп.

- 1) Языки (или их диалекты), которые заснули на территории России, но продолжают свое существование в других странах. К таким языкам можно отнести алеутский (беринговский диалект), караимский (крымский диалект) и крымчакский языки.
- 2) Заснувшие языки, которые существовали только на территории России: бабинский саамский, камасинский, керекский, ороцкий, сиреницкий, сойотский, югский и южнохантыйский языки.
- 3) Языки, которые продолжили свое существование в других странах, поскольку их носители переселились с территории России. К таким языкам относятся убыхский и язык эскимосов Берингова пролива.
- 4) Отдельный случай представляет собой сахалинско-айнский язык.

Рассмотрим эти языки подробнее.

Алеутский язык принадлежит к эскимосско-алеутской языковой семье, в составе которой образует отдельную алеутскую ветвь. Исторически его носители проживали на Алеутских островах, которые сегодня входят в состав американского штата Аляска. В первой половине XIX века колониальная Российско-американская компания, созданная царским правительством, чтобы контролировать этот архипелаг, переселила часть алеутов на Командорские острова, расположенные к востоку от побережья Камчатки. На Командорах развился особый диалект алеутского языка — беринговский, названный так по крупнейшему в архипелаге острову Беринга. Последняя полноценная носительница, Вера Терентьевна Тимошенко, умерла в селе Никольское на этом острове в марте 2021 года [Головко 2021]. Остальные алеуты могут знать отдельные слова и выражения из этнического языка, но в быту общаются исключительно по-русски. Тем не менее, в школе села Никольского дети изучают алеутский как иностранный — уроки ведет Елена Солованюк, которая выучила этот язык по книгам во взрослом возрасте. Всего в мире насчитывается около 460 носителей алеутского, большинство из которых проживает на островах Прибылова в США (другой диалект).

Караимский язык относится к тюркской языковой семье. В караимском выделяют три диалекта: тракайский (диалект литовских караимов), галичский (диалект караимов Западной Украины) и крымский. В то время как два первых диалекта еще живы, диалект караимов Крыма фактически прекратил свое существование в 2000-х гг. Мария Полински вела работу с крымскими караимами в 1980-х годах, в настоящее время

полноценных носителей не осталось. По данным Н. В. Кропотовой, в России сейчас есть около пяти носителей, которые знают счет и отдельные фразы.

Крымчакский язык также относится к тюркской языковой семье. Одна из последних носительниц крымчакского — Виктория Багинская (девичья фамилия — Гурджи) — умерла в январе 2012 года в Краснодаре. Последним свободно говорящим в Крыму был автор учебника крымчакского языка Давид Ильич Реби, умерший в Симферополе в апреле 2019 года. По сведениям Д. Т. Пирковой, в Германии и Израиле остаются репатрианты, знающие крымчакский язык.

Бабинский саамский, или аккала саамский, принадлежит к уральской семье языков (о выделении бабинского саамского как самостоятельного языка см. раздел 4.4.7.5). В работе [Rantala, Sergina 2009] сообщается, что последний носитель бабинского саамского умер в 2003 году. В [Шеллер 2010] содержится информация о том, что осталось два человека в возрасте старше 70 лет, которые имеют определенные знания бабинского саамского языка, но которые на нем не говорят, то есть являются пассивными носителями этого языка. Они проживают вне компактных поселений бабинских саамов. Предполагается, что в селе Ёна Ковдорского района Мурманской области, которое является компактным поселением бабинских саамов, могут быть люди, которые слышали бабинский саамский в детстве и тем самым могут иметь определенные пассивные знания бабинского языка. В своей статье Э. Шеллер сообщает, что в селе Ёна (по крайней мере, по состоянию на 2010 год) организуются курсы кильдинского языка; среди бабинских саамов, посещающих эти курсы, есть люди, которые надеются, что, выучив кильдинский, они смогут перейти на бабинский.

Камасинский язык также относится к уральской семье. Последняя носительница камасинского языка, Клавдия Захаровна Плотникова, жила в селе Абалаково Красноярского края и умерла в 1989 году. В социальных сетях Facebook и ВКонтакте³⁷ ведутся группы, посвященные камасинскому языку. Создателем этих групп также ведется канал на YouTube, где публикуются тексты песен, переведенных на камасинский язык³⁸.

Керекский язык принадлежит к чукотско-камчатской семье. По сведениям лингвистов, в 2005 году умер последний носитель керекского, хотя в данных

³⁷ См. <https://www.facebook.com/groups/kamassian/> и <https://vk.com/kamassian>.

³⁸ См. https://www.youtube.com/channel/UChACHjoKu73_ZDxqAqvncpA/about.

Всероссийской переписи 2010 года 10 человек указали владение керекским [Муравьева 2018].

Как сообщается в [Перехвальская 2016], на данный момент можно считать, что носителей ороцкого языка, относящегося к тунгусо-маньчжурской семье, не осталось. Ассоциация коренных малочисленных народов Хабаровского края заявила, что последний носитель ороцкого языка умер в 2008 году. Анатолий Филиппович Намунка, о котором говорится в документе, владел своим этническим языком далеко не свободно. 5 августа 2010 года собравшиеся в помещении Ассоциации наиболее пожилые ороочи сообща вспомнили около 20 отдельных слов, но не смогли восстановить полностью счет от 1 до 10.

Сиреникский язык относится к эскимосской ветви эскимосско-алеутской языковой семьи. Носители сиреникского языка обитали в прошлом на юго-восточном побережье Чукотки в поселке Сиреники и в нескольких небольших прилегающих поселках к западу от него. В 1997 году умерла последняя носительница языка — Валентина Выйе [Вахтин 2000]. Полиглот Егор Антонов из Нижегородской области самостоятельно решил выучить сиреникский язык и является новым носителем этого языка (на 2019 год) [Википедия].

Сойотский язык принадлежит тюркской семье. На нем изначально говорили сойоты, являющиеся потомками саянских самодийцев; впоследствии сойоты были подвержены тюркизации, а затем — влиянию бурятского языка. Лингвистически наиболее близок сойотскому языку тофаларский. В [Рассадин 2012] содержатся свидетельства того, что еще в 1970-е гг. и даже в 1990-е гг. можно было услышать сойотский язык. В работе сообщается, что разработаны и принимаются меры по возрождению национальной культуры и промысла, а также языка. Ученицей В. И. Рассадина Сэсэгмой Гармаевой ведется преподавание сойотского языка в школе села Сорок³⁹.

Югский язык входит в енисейскую языковую семью. Долгое время рассматривался как диалект кетского языка. По историческим данным можно предполагать, что югский был распространен от г. Енисейска до р. Дупчес, впадающей в Енисей недалеко от с. Ворогова [Вернер 1997]. В настоящее время сохранилось лишь несколько семей югов в селе Ярцево и селе Ворогово на Енисее, которые родным языком уже не пользуются. Полевые исследования, проводившиеся в начале 1980-х, показали,

³⁹ См. <https://www.nordgold.com/ru/soyots/>.

что тогда югским языком владело 2 человека, а обследование В. П. Кривоногова в 1992 году выявило, что уже к тому времени носителей языка не осталось [Казакевич 2003].

Южнохантыйский язык относится к уральской семье (о выделении южнохантыйского как самостоятельного языка см. раздел 4.4.7.7). На южнохантыйском перестали говорить приблизительно в середине XX в. [А. Ю. Фильченко, л. с.], точную дату исчезновения южнохантыйского языка определить уже невозможно.

Убыхский язык принадлежит к абхазо-адыгской семье. Был распространен на территории нынешних Лазаревского, Центрального и Хостинского районов Сочи и исчез в России из-за того, что после окончания Кавказской войны (1864 г.) убыхи были выселены правительством России в Османскую Турцию. Последний носитель этого языка скончался в Турции в 1992 году.

Название “язык эскимосов Берингова пролива” принято в отечественном языкознании и обусловлено ареалом распространения. Язык принадлежит к эскимосско-алеутской языковой семье. В 1948 г. носители из поселения Имаклик были переселены в Наукан, из которого в 1958 г. жители были выселены за рубеж; в настоящее время на территории США остаются говорящие носители.

Среди остальных групп заснувших языков особый случай представляет собой сахалинско-айнский язык, так как его носители были разделены на две группы: после Второй мировой войны большая часть носителей была переселена в Японию, однако часть носителей осталась жить на Сахалине. Мы выделяем сахалинско-айнский как самостоятельный язык, так как в [Алпатов 1997] сообщается, что совпадение в базовой лексике по списку Сводеша между сахалинскими диалектами и диалектами юга Хоккайдо находится в пределах 68-76%⁴⁰. В [Алпатов 2004] сообщается, что последний носитель сахалинско-айнского, проживавший в СССР, умер в городе Анива (о. Сахалин) в 1975 году. Последний носитель сахалинско-айнского, проживавший в Японии, скончался в 1994 году.

7. Другие проблемы

Как правило, при перечислении языков России используется списочное представление (как это сделано и в нашем предварительном списке в Приложении 1),

⁴⁰ Отметим, однако, что подсчет нормализованного расстояния Левенштейна между списками айнских диалектов, содержащихся в базе данных ASJP, не дает превышения порога в 0,51. По данной метрике это говорит о единстве айнского языка.

однако возможно также представление в виде генетической классификации языков. Тем не менее, при таком представлении возникают некоторые проблемы. Во-первых, для некоторой группы языков может существовать несколько альтернативных классификаций, ср., например, различные классификации для тюркских языков Г. Х. Ахатова, Н. А. Баскакова, В. А. Богородицкого, В. В. Радлова и др. (алабугатско-ногайский, юртовско-ногайский, карагашский [Арсланов 1980; Булгарова и др. 2020], сибирскотатарский язык [Ахатов 1963; Исхакова 1973; Закиев 1995]). Во-вторых, проблему будут представлять смешанные языки⁴¹ (например, алеутско-медновский [Головко 2011]), которые “выпадают” из традиционной классификации. Кроме того, существуют различные гипотезы по объединению языковых семей в более крупные единицы, принимаемые не всеми учеными (ср. алтайскую гипотезу, согласно которой тюркские, монгольские и тунгусо-маньчжурские языки имеют общее происхождение от праалтайского языка).

При составлении списка языков России представляется важным учитывать уровень диалектов, ведь это способствует привлечению внимания к их состоянию и может помочь сохранению всех вариантов языка. Однако это также вызывает ряд трудностей. Во-первых, у исследователей может различаться мнение по поводу диалектного членения того или иного языка, см., например, обзор диалектного членения кильдинского саамского в [Эрштадт 2014: 12-17]. Во-вторых, классификация диалектов может быть спорной: например, аккинско-орстхойское наречие традиционно причисляют к чеченскому языку, но в работе [Коряков 2006] оно называется переходным между чеченским и ингушским языками. В-третьих, диалекты какого-то языка, представленные на удаленных территориях, могут разойтись из-за влияния соседних языков: к примеру, диалекты шорского языка в настоящий момент относят к разным подразделениям хакасской подгруппы [Баскаков 1952: 123]. В-четвертых, для экстратерриториальных языков также надо будет указывать, носители каких диалектов представлены в России (например, большинство финнов в России владеют ингерманландским диалектом финского языка [Шлыгина 1994]), а это требует во многих случаях дополнительных исследований.

⁴¹ Смешанные языки появляются в ситуации двуязычия, когда носители достаточно хорошо знают оба языка, чтобы из их элементов сконструировать новый язык (не путать со смешением кодов).

8. Перспективы дальнейшей работы

Как уже было сказано выше, список языков может меняться со временем из-за появления новых данных или в зависимости от задач исследователя. И наш предварительный список — не исключение. Тем не менее, он мог бы послужить основой для следующей всероссийской переписи населения, наряду с рекомендациями ИЭА РАН.

Необходимо детальное изучение тех идиомов, которые не выделялись ранее в отдельные языки, так как, как правило, различной социолингвистической информации в литературе бывает недостаточно.

Как было упомянуто, нам представляется важным спуститься на уровень диалектов, изучив их перечень и состояние. И тогда список может пополниться теми идиомами, которые находятся в угрожаемом состоянии, но на момент составления текущего списка у нас такой информации не было.

В качестве будущего исследования также полезно было бы составить более полный перечень методик измерения взаимопонятности и апробировать их на материале языков России — в отечественной лингвистике таких исследований крайне мало, хотя нарушенная взаимопонятность некоторого идиома с другими диалектами и особенно с литературной нормой негативно влияет на его функционирование.

9. Заключение

Данный документ посвящен проблеме составления списка языков России. Нами были рассмотрены перечни языков России, содержащиеся в различных энциклопедиях и каталогах языков мира. Лишь в редких случаях авторы этих изданий и ресурсов указывают критерии, по которым они были составлены. Целенаправленная работа по составлению списка языков России и выработке критериев, обосновывающих включение различных языков в список, ведется в Институте языкознания РАН под руководством Ю. Б. Корякова. Однако в документах, доступных на сайте этого проекта, также не всегда очевидно, почему один язык входит в список, а другой — нет.

В настоящем документе рассматриваются основные проблемы, с которыми можно столкнуться при составлении списка языков России.

Наиболее значительной проблемой (причем для составления списка языков в целом, а не только списка языков России) является определение границы между языком и диалектом. В работе рассматриваются основные критерии, связанные с этой

проблемой. Акцент делается на лингвистических критериях, а также на критерии взаимопонятности. Отметим, что исследований, направленных на выявление степени взаимопонятности между идиомами, на материале языков России проведено крайне мало. Хотелось бы надеяться, что составленный нами краткий перечень современных методик измерения взаимопонятности может подтолкнуть исследователей к развитию данного направления в отечественной лингвистике. Среди языков, распространенных на территории России, можно выделить множество идиомов, которые подпадают под проблему “язык или диалект”. Каждый такой идиом анализируется с точки зрения подобранных нами критериев. На основании проведенного анализа выносится решение, следует считать данный идиом языком или нет. Составленный список проблемных идиомов представляется нам существенным заделом для будущих исследований, касающихся как составления списка языков России, так и важной для лингвистической науки проблемы разграничения языка и диалекта.

Другая важная проблема касается языков, носители которых проживают преимущественно за пределами России. Такие языки мы условно называем экстратерриториальными. Для этих языков нами также был составлен перечень критериев, которые затем были к ним применены. Для каждого экстратерриториального языка, вошедшего в предварительный список языков России, в данном документе содержатся краткие обоснования для его включения в список.

Проблему для составления списка языков России представляет также учет языков, признанных исчезнувшими на территории нашей страны. Так, имеющиеся социолингвистические данные не всегда дают представление о том, действительно ли в России не осталось больше носителей предположительно исчезнувшего идиома. Кроме того, важно учитывать нацеленность сообщества на восстановление языка. Благодаря деятельности языковых активистов, языкового сообщества и лингвистов языки могут возвращаться к жизни. Именно поэтому мы предпочитаем использовать термин “заснувшие языки”. Нами был составлен список языков, заснувших на территории России с середины XX века. Некоторые из этих языков все еще имеют пассивных носителей, с которыми возможно было бы выстроить стратегию ревитализации их языка.

При составлении списка языков России мы также обратили внимание и на другие проблемы. Одна группа проблем возникает в случае, когда языки представляются не в виде списка, а в виде генетической классификации. Другая группа проблем связана с учетом уровня диалектов.

В заключение хотелось бы еще раз отметить, что любые списки языков России не являются окончательными. Необходимо регулярное обновление списочного состава с учетом данных, поступающих от специалистов и исследователей.

10. Список использованной литературы

Закон Российской Федерации «О языках народов Российской Федерации» от 25 октября 1999 г. № 1807-1 // СПС «Кодекс». URL: <https://docs.cntd.ru/document/9003298> (дата обращения: 19.11.2021).

Закон Республики Алтай «О языках народов, проживающих на территории Республики Алтай» от 03 марта 1993 г. №9-6 // СПС «Кодекс». URL: <https://docs.cntd.ru/document/802008526> (дата обращения: 19.11.2021).

Закон Республики Карелия «О государственной поддержке карельского, вепсского и финского языков в республике Карелия» от 19 марта 2004 г. №759-ЗРК // СПС «Кодекс». URL: <https://docs.cntd.ru/document/919316706> (дата обращения: 19.11.2021).

Конституция Республики Крым от 11 апреля 2014 г. // СПС «Кодекс». URL: <https://docs.cntd.ru/document/413901001> (дата обращения: 19.11.2021).

Постановление Правительства Российской Федерации «О едином перечне коренных малочисленных народов Российской Федерации» от 24 марта 2000 г. №255 // СПС «Кодекс». URL: <https://docs.cntd.ru/document/901757631> (дата обращения: 19.11.2021).

Постановление Государственного совета Республики Дагестан «О коренных малочисленных народах Республики Дагестан» от 18 октября 2000 г. №191 // СПС «Кодекс». URL: <https://docs.cntd.ru/document/473101930> (дата обращения: 19.11.2021).

Постановление Правительства Республики Карелия «Об утверждении алфавитов карельского и вепсского языков» от 16 марта 2007 г. №37-П // СПС «Кодекс». URL: <https://docs.cntd.ru/document/919323769> (дата обращения: 19.11.2021).

Рекомендации Совета при Президенте Российской Федерации по развитию гражданского общества и правам человека по итогам 16-го выездного специального заседания в Тверской области 27-29 июля 2016 г. // URL: <http://president-sovet.ru/files/92/d6/92d6d4d6d01775e7738e6e18d69b3aad.pdf> (дата обращения: 19.11.2021).

Рекомендации Совета при Президенте Российской Федерации по развитию гражданского общества и правам человека по итогам 26-го выездного специального заседания в Новосибирской области 14-17 августа 2018 г. // URL: <http://president-sovet.ru/files/0c/e9/0ce9231c17646ba50cfb4b604d23c216.pdf> (дата обращения: 19.11.2021).

Устав Еврейской автономной области от 8 октября 1997 г. №40-ОЗ // СПС «Кодекс». URL: <https://docs.cntd.ru/document/804930931> (дата обращения: 15.12.2021).

Valsts valodas likums. 21.12.1999 // URL: <https://likumi.lv/ta/id/14740-valsts-valodas-likums> (дата обращения: 15.12.2021).

Zákon Národnej rady Slovenskej republiky č. 184/1999 Z.z. o používaní jazykov národnostných menšín z 10. júla 1999 // URL: <https://www.epi.sk/zz/1999-184> (дата обращения: 19.11.2021).

Агранат Т. Б. Водский язык // Большая российская энциклопедия : [сайт]. URL: <https://bigenc.ru/linguistics/text/5200027> (дата обращения: 11.10.2021).

Агранат Т. Б. Язык, культура и традиции печорских сето // Вестник МГЛУ. Гуманитарные науки. 2019. Вып. 4 (820). С. 219-226.

Агранат Т. Б. Сето в Сибири // Материалы Всероссийской научной конференции с международным участием «Бубриховские чтения: задокументированное народное слово» (Петрозаводск, 27-28 октября 2020 г.) / отв. ред. И. П. Новак. Петрозаводск: Издательство ПетрГУ, 2020. С. 38-41.

Агранат Т. Б. Язык сето // ПостНаука : [сайт]. URL: <https://postnauka.ru/longreads/156169> (дата обращения: 19.11.2021).

Алексеев М. Е. Языки народов России // Большая российская энциклопедия : [сайт]. URL: <https://bigenc.ru/text/3040752> (дата обращения: 19.11.2021).

Алексеев М. Е. Нахско-дагестанские языки // Языки мира. Кавказские языки. — М.: Academia, 1999. — С. 156—165.

Алексеев М. Е. Ахвахский язык // Язык и общество. Энциклопедия. М.: Издательский центр «Азбуковник», 2016а. С. 77-78.

Алексеев М. Е. Чамалинский язык // Язык и общество. Энциклопедия. М.: Издательский центр «Азбуковник», 2016б. С. 551-552.

Алексеев Ф. Г. Татарско-казахские ногайцы: языковая идентичность астраханских карагашей // Первая конференция по уральским, алтайским и палеоазиатским языкам. Памяти А. П. Володина. Тезисы докладов международной научной конференции, Санкт-Петербург 6-8 декабря 2018 г. СПб.: ИЛИ РАН, 2018. С. 9-11.

Алпатов В. М. Айнский язык // Языки мира. Палеоазиатские языки. М.: Издательство «Индрик», 1996. С. 126-138.

Андронов А. В. Латгальский литературный язык в свете лексикостатистики // Латгальской грамматике Антоны Скрынды 100. 1-ая международная конференция по латгалистике (Санкт-Петербург, 19-20 сентября 2008 г.). Тезисы докладов / ред. А. В. Андронов, Л. Лейкума, И. Шуплинска. СПб.: Факультет филологии и искусств Санкт-Петербургского государственного университета, 2008.

Арсланов Л. Ш. Формирование и развитие островных языков и диалектов (на материале тюркских языков и диалектов Волгоградской, Астраханской областей, Ставропольского края и Калмыцкой АССР). Дисс. на соиск. уч. степ. д. филол. наук. Елабуга: Елабужский государственный педагогический институт, 1980.

Арсланов Л. Ш. Язык карагашей-ногайцев Астраханской области. Набережные Челны: КГПС, 1991.

Арсланов Л. Ш. Алабугатских татар язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 1. С. 57-63.

Арсланов Л. Ш. Юртовских татар язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 3. С. 507-513.

Ахатов Г. Х. Диалект западносибирских татар. Уфа: Башкнигоиздат, 1963.

Баскаков Н. А. К вопросу о классификации тюркских языков // Известия АН СССР. Отделение литературы и языка. 1952. Т. XI. Вып. 2. С. 121-134.

Беликов В. И. Многоязычный Дагестан: лингводемографический комментарий к Всероссийской переписи населения 2010 // Малые языки в большой лингвистике. Сборник трудов конференции 2020 / ред. Кс. П. Семёнова. М.: «Буки Веди», 2020. С. 16-24.

Бирин В. Н., Такала И. Р. Финны // Народы России. Энциклопедия / гл. ред. В. А. Тишков. М.: Научное издательство «Большая российская энциклопедия», 1994. С. 370-372.

Бокарев Е. А. Краткие сведения о языках Дагестана. Махачкала, 1949.

Брейдак А. Б. Латгальский язык // Языки мира. Балтийские языки. М.: Academia, 2006. С. 193-213.

Брук С. И., Смирнова Т. Б. Немцы // Народы России. Энциклопедия / гл. ред. В. А. Тишков. М.: Научное издательство «Большая российская энциклопедия», 1994. С. 246-249.

Булгарова М. А. Ногайский язык // ПостНаука : [сайт]. URL: <https://postnauka.ru/longreads/155734> (дата обращения 27.07.2021).

Вахтин Н. Б. Язык сиреникских эскимосов: Тексты, грамматические и словарные материалы. Munchen: Lincom-Europa, 2000.

Вахтин Н. Б. Эскимосский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 593-596.

Вернер Г. К. Югский язык // Языки мира. Палеоазиатские языки. М.: Издательство «Индрик», 1997. С. 187-195.

Влахов А. В., Нагурная С. В., Новак И. П., Пашкова Т. В. Карельский язык // ПостНаука : [сайт]. URL: <https://postnauka.ru/longreads/155709> (дата обращения 19.11.2021).

Вольнец А. «К выселению приступить немедленно...» История депортации корейцев Дальнего Востока // ТАСС. Дальний Восток : [сайт]. URL: <https://dv.land/history/k-vyseleniyu-pristupit-nemedlenno> (дата обращения 16.11.2021).

Гасанова У. У. Лексика и словообразование хайдакского диалекта даргинского языка. Махачкала, 2011.

Герцен Е. Светлана Ким: корейское телевидение — мечта, ставшая явью // Сахалин Р. S. : [сайт]. URL: <https://sakhalin-ps.ru/lyudi-ostrovov/svetlana-kim-koreyskoe-televeshchanie--mechta-stavshaya-yavyu/> (дата обращения 16.11.2021).

Головко Е. В. Медновско-алеутский язык // Большая российская энциклопедия : [сайт]. URL: <https://bigenc.ru/linguistics/text/2198246> (дата обращения 19.11.2021).

Головко Е. В. Алеутский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 54-58.

Головко Е. В. Минус алеутский: каким был язык, который мы потеряли // Онлайн-заседание отдела языков народов России ИЛИ РАН 19 марта 2020 г.

Грачева Ф. Т. Марийский горный язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 2. С. 274-283.

Грюнберг А. Л. Татский язык // Языки мира. Иранские языки. I. Юго-западные иранские языки. М.: Индрик, 1997. С. 141-153.

Дмитриев С. Д., Дмитриева В. М., Тужаров Г. М. Маре букварь. Йошкар-Ола: Марийское книжное издательство, 1995.

Ельяшевич В. А. Проблема классификации разговорного языка крымских караимов и перспективы ее решения // Научный вестник Крыма. 2020. Вып. 2 (25).

Жукова А. Н. Корякский язык // Языки мира. Палеоазиатские языки. М.: Издательство «Индрик», 1996. С. 39-53.

Закиев М. З., Ганиев Ф. А., Зиннатуллина К. К. (ред.). Татарская грамматика. В трех томах. Том I. Введение. Фонетика. Фонология. Просодика. Графика и орфография. Орфоэпия. Основные понятия грамматики. Морфемика. Морфонология. Словообразование. Казань: Татарское книжное издательство, 1995.

Ибрагимов А. Г. Гельмецкий диалект цахурского языка: диссертация ... кандидата филологических наук - Махачкала, 2011.- 151 с.

Исрафилов Н. Р. Фитинский говор агульского языка: автореферат дис. ... кандидата филологических наук - Махачкала, 2007. - 21 с.

Исхакова С. М. Древнетюркские элементы в народноразговорном языке западносибирских татар // Советская тюркология. 1973. №6. С. 36-39.

Ишмухамбетов Р. В. К вопросу о “плавающей” идентичности у астраханских ногайцев // Астраханские краеведческие чтения. Выпуск XII: сборник статей / ред. А. А. Курапов, А. Н. Алиева. Астрахань: Издатель: Сорокин Роман Васильевич, 2020. С. 374-381.

Казакевич О. А. Югский язык // Письменные языки мира: Языки Российской Федерации. М.: Academia, 2003. Т. 2. С. 732-738.

Казакевич О. А. Селькупский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 418-425.

Казакевич О. А., Сердобольская Н. В., Халилова З. М., Будянская Е. М., Евстигнеева А. П., Мордашова Д. Д., Покровская С. В., Поливанов К. К., Ренковская Е. А., Шейфер К. О. Шкалы витальности языков и проблемы их использования. Третья научная конференция “Малые языки в большой лингвистике 2021”.

Кашкин Е. В., Мордашова Д. Д. Горномарийский язык // ПостНаука : [сайт]. URL: <https://postnauka.ru/longreads/155718> (дата обращения 19.11.2021).

Калимуллин Т. Р. Как проходила перепись населения в 2002 г.: наблюдения участников // Экономическая социология. 2003. Т. 4. №1. С. 147-157.

Калинина Е. Ю., Оскольская С. А. Нанайский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 293-296.

Керт Г. М. Саамский язык // Языки народов СССР. В пяти томах. М.: Наука, 1966. Т. 3. С. 155-171.

Ким Г. Н. Материковские и сахалинские корейцы: различия и сходства // Koreans.kz : [сайт]. URL: <https://koreans.kz/news/materikovskie-i-sahalinskie-koreycy-razlichiya-i-shodstva.html> (дата обращения 16.11.2021).

Ковалева С., Кондрашкина Е. А. Карельский язык // Язык и общество. Энциклопедия. М.: Азбуковник, 2016. С. 203-207.

Коростелев А. Д. К вопросу о соотношении языка горных и луговых марийцев // Финно-угроведение. 2003. № 2. С. 31-39.

Коряков Ю. Б. Атлас кавказских языков. М.: Пилигрим, 2006.

Коряков Ю. Б. Лексикостатистическая классификация даргинских языков. Московский семинар по нахско-дагестанским языкам под руководством Н. Р. Сумбатовой 30 октября 2012 г.

Коряков Ю. Б. Проблема «язык или диалект» и попытка лексикостатистического подхода // Вопросы языкознания. 2017. № 6. С. 79-101.

Коряков Ю.Б. Проблема язык или диалект и самодийские языки // Урало-алтайские исследования. 2018а. № 4 (31). С. 156-217.

Коряков Ю. Б. Амшенский армянский: история изучения и современное состояние // Малые языки в большой лингвистике. Сборник трудов конференции 2017 / ред. Кс. П. Семёнова. М.: «Буки Веди», 2018б. С. 100-106.

Коряков Ю. Б. Принципы составления списка языков России // 3-е заседание Дискуссионно-аналитического клуба по вопросам языковой политики 24 ноября 2020 г.

Коряков Ю. Б. Даргинские языки и их классификация // Дурхъаси хазна. Сборник статей к 60-летию Р. О. Муталова / ред. Т. А. Майсак, Н. Р. Сумбатова, Я. Г. Тестелец. М.: Буки Веди, 2021. С. 139-154.

Коряков Ю. Б., Майсак Т. А. Систематика языков мира и базы данных в интернете // Труды Международного семинара «Диалог 2001» по компьютерной лингвистике и ее приложениям. Т. 2. М., Аксаково, 2001. С. 215-229.

Коряков Ю. Б., Сумбатова Н. Р. Даргинские языки // Большая российская энциклопедия, Т. 8. М.: БРЭ, 2007.

Коряков Ю. Б. (рук.), Агранат Т. Б., Горячева М. А., Дыбо А. В., Казакевич О. А., Кибрик А. А., Ханина О. В., Шлуинский А. Б. Проект «Языки России». URL: <http://jazykirf.iling-ran.ru/> (дата обращения 19.11.2021).

Кумахов М. А. (гл. ред.). Адыгская (черкесская) энциклопедия. М.: Фонд им. Б. Х. Акбашева, 2006.

Курилов Г. Н. Лексикология современного юкагирского языка: развитие лексики и роль в нем якутского языка. Новосибирск: Наука, 2003.

Кусаева З. К., Сатцаев Э. Б., Таказов Ф. М. Джалган и джалганцы: синтез персидского и аланского компонентов (по материалам фольклорно-этнографической и лингвистической экспедиции 14-18 июня 2016 г.) // Известия Северо-Осетинского Института гуманитарных и социальных исследований. 2016. Вып. 21 (60). С. 96-104.

Магомедов А. А. Кубачинский язык. (Исследования и тексты). Тбилиси, 1963.

Мальцева А. А. Алюторский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 62-69.

Межекова Н. Н. Шорский диалект // Диалекты хакасского языка. Абакан: Хакасия, 1973. С. 60-75.

Меновщиков Г. А. Язык сиреникских эскимосов. Фонетика, очерк морфологии, тексты и словарь. М.–Л.: Наука, 1964.

Муравьева И. А. Керекский язык // Большая российская энциклопедия : [сайт]. URL: <https://bigenc.ru/linguistics/text/5229835> (дата обращения 13.11.2021).

Мусаев М.-С.М. История словоизменительных категорий даргинского языка. АДД. М. 1989.

Муталов Р. О. Классификация даргинских языков и диалектов // 10-е заседание Дискуссионно-аналитического клуба по вопросам языковой политики 23 сентября 2021 г.

Намруева Л. В. Как калмыки знают свой язык // Социологические исследования. 2010. №4. С. 138-141.

Немировский А.А. Новые лингвистические результаты М. А. Живлова и подтверждение ымыяхтахского соотнесения (пра)юкагиров // Фольклор палео-азиатских народов: матер. II Междунар. науч. конф. (г. Якутск, 21–25 ноября 2016 г.). Якутск: РИО медиа-холдинга, 2017.

Николаева И.А., Хелимский Е.А. Юкагирский язык // Языки мира. Палеоазиатские языки. М.: «Индрик», 1996. С. 155-168.

Норманская Ю. В. Коми-язьвинский — диалект коми-пермяцкого или отдельный язык? // Ежегодник финно-угорских исследований. 2020. Т. 14. №4. С. 630-643.

Оскольская С. А. Аспект в нанайском языке. Дисс. на соиск. уч. степ. к. филол. наук. СПб.: ИЛИ РАН, 2017.

- Панфилов В. З. Нивхский язык // Языки народов СССР. В пяти томах. М.: Издательство «Наука», 1966. Т. 5. С. 408-434.
- Паршакова А. Л. Коми-язьвинский букварь. Пермь: Пермское книжное издательство, 2003.
- Паршакова А. Л. Лөддьөтан книга. Книга для чтения: хрестоматия на коми-язьвинском языке. Пермь: Пермский государственный педагогический университет, 2008.
- Перехвальская Е. В. Статус идиома и этническое самосознание // Малые языки Евразии: социолингвистический аспект. 1997. С. 5-12.
- Перехвальская Е. В. Язык или диалект: современное состояние проблемы // Актуальные проблемы диалектологии языков народов России. Материалы XII региональной конференции (Уфа, 27–28 ноября 2012 г.). 2012. С. 117-120.
- Перехвальская Е. В. Ороцкий язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 339-342.
- Поливанов К. К., Казакевич О. А., Сердобольская Н. В., Халилова З. М., Будянская Е. М., Евстигнеева А. П., Мищенко К. О., Мордашова Д. Д., Покровская С. В., Ренковская Е. А. Информационный ресурс «Малые языки России»: задачи и специфика // Малые языки в большой лингвистике. Сборник трудов конференции 2020 / ред. Кс. П. Семёнова. М.: «Буки Веди», 2020. С. 186-193.
- Пунжина А. В. (сост.). Слушаю карельский говор. Образцы речи дёржанских и валдайских карел. Петрозаводск: «Периодика», 2001.
- Рассадин В. И. Язык сойотов Бурятии. Элиста, 2012.
- Рогава Г. В. Абхазо-адыгские языки. Введение // Языки народов СССР. В пяти томах. М.: Издательство «Наука», 1966. Т. 4. С. 95-100.
- Рягоев В. Д. Карельский язык // Языки мира. Уральские языки. М.: «Наука», 1993. С. 63-75.
- Сагидуллин М. А. Фонетика и графика современного сибирскотатарского языка. Тюмень: Искер, 2008.
- Сагидуллин М. А. Русско--сибирскотатарский словарь (Урысца -себертатарца сүслек). Тюмень : Мандр и К^а, 2010.
- Сатанова Д. С. Язык ногайцев-карагашей (астраханско-ногайский язык): краткий обзор социолингвистической ситуации // Малые языки. 2015. №1. С. 17-34.
- Селютина И. Я., Уртегешев Н. С., Эсенбаева Г. А., Добринина А. А., Рыжикова Т. Р. Атлас консонантных артикуляций в тюркских языках народов Сибири. Новосибирск: Ин-т филологии СО РАН, 2013.
- Суник О. П. Кур-урмийский диалект. Материалы и исследования по нанайскому языку. Л.: Государственное учебно-педагогическое издательство Министерства просвещения РСФСР, Ленинградское отделение, 1958.

Старостин Г. С. Языки Африки. Опыт построения лексикостатистической классификации. М., 2012.

Сызранов А. В. Этносы и этнические группы Астраханской области. Астрахань: ОмЦ НК, 2008.

Темирбулатова С. М. Хайдакский диалект даргинского языка. Махачкала: ДНЦ РАН, 2004.

Терещенко Н. М. Ненецкий язык // Языки народов СССР. В пяти томах. М.: Издательство «Наука», 1966. Т. 3. С. 376-395.

Трофименко О. А. Фонетические особенности диалекта корё мар (на примере языка корейцев Приморского края) // Вестник Орловского государственного университета. Серия: Новые гуманитарные исследования. 2012. №7. С. 231-233.

Урманчиева А. Ю. Фрагмент описания глагольной морфологии энецкого языка: соотношение синхронии и диахронии // Языки мира. Типология. Уралистика. Памяти Т. Ждановой. Статьи и воспоминания / сост. Плунгян В. А., Урманчиева А. Ю. М., 2002. С. 573-591.

Уртегешев Н. С. Калмакский язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 186-189.

Федотова И. В. Диалектное членение мансийского языка в диахронии по данным базисной лексики // Вестник урведования. 2021. Т. 11. №2. С. 338-346.

Федюнева Г. В. О прибалтийско-финском компоненте в коми языке // Известия Уральского государственного университета. 2008. №55. С. 172-180.

Феоктистов А. П. Мордовский-эрзя язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 2. С. 306-315.

Хайдаков С. М. Даргинский и мегебский языки: Принципы словоизменения. М., 1985.

Хастян А. А. Армянские субобщности на Северном Кавказе: факторы социокультурной адаптации. Дисс. на соиск. уч. степ. к. социол. наук. Ростов-на-Дону: Ростовский государственный педагогический университет, 2001.

Хелимский Е. А. Селькупский язык // Языки народов России. Красная книга. Энциклопедический словарь-справочник / гл. ред. В. П. Нерознак. М.: «Academia», 2002а. С. 160-164.

Хелимский Е. А. Хантыйский язык // Языки народов России. Красная книга. Энциклопедический словарь-справочник / гл. ред. В. П. Нерознак. М.: «Academia», 2002б. С. 211-215.

Хомякова Д. 2021. Непризнанный народ // Наука в Сибири : [сайт] URL: <https://www.sbras.info/articles/simply/nepriznannyyi-narod> (дата обращения 19.11.2021).

Хонти Л. Хантыйский язык // Языки мира. Уральские языки. М.: «Наука», 1993. С. 301-319.

Хонт Л., Николаева И. А., Каксин А. Д. Хантыйский язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 3. С. 301-311.

Цинциус В. И. Солонский язык // Языки мира. Монгольские языки. Тунгусо-маньчжурские языки. Японский язык. Корейский язык. М.: Издательство «Индрик», 1997.

Чирикба В. А. Абазинский язык // Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. М.: Наука, 2001. Т. 1. С. 1-8.

Шабыков В. И., Кудрявцева Р. А., Зорина З. Г. Статус горномарийского языка в современном обществе // Социодинамика. 2020. №4. С. 74-84.

Шеллер Э. Ситуация саамских языков в России // Наука и бизнес на Мурмане. Мурманск: Мурманское областное книжное издательство. 2010. №2 (69). С. 15-27.

Шкляев Г. К. О бесермянах: опыт этностатистического обследования // О бесермянах. Сборник статей / отв. ред. Г. К. Шкляев. Ижевск: Удмуртский институт истории, языка и литературы УрО РАН, 1997. С. 110-120.

Шлуинский А. Б., Ханина О. В. Энецкий язык // Язык и общество. Энциклопедия. М.: «Азбуковник», 2016. С. 588-593.

Эдельман Д. И. К проблеме «язык или диалект» в условиях отсутствия письменности // Теоретические основы классификации языков мира / ред. В. Н. Ярцева. М.: Наука, 1980. С. 127-147.

Эрштадт А. М. Лексика традиционных хозяйственных занятий кольских саамов (на материале кильдинского диалекта саамского языка. Дисс. на соиск. уч. степ. к. филол. наук. М.: Мурманский государственный гуманитарный университет, 2014.

Юзыкайн Э. Некоторые аспекты информационной среды марийцев. Магистерская диссертация. Таллинн: Таллиннский педагогический университет, 2000.

Яхонтов С. Е. Оценка степени близости родственных языков // Теоретические основы классификации языков мира / ред. В. Н. Ярцева. М.: Наука, 1980. С. 148-157.

Курды Адыгеи / гл. ред. А. Н. Соколова; ред. А. Ю. Шадже, З. А. Жаде. Майкоп: Качество, 2015.

Народы России: Атлас культур и религий / отв. ред. В. А. Тишков, А. В. Журавский, О. Е. Казьмина. М.: ИПЦ «Дизайн. Информация. Картография», 2008.

Письменные языки мира: Российская Федерация. Социолингвистическая энциклопедия. Кн. 1. / отв. ред. В. М. Солнцев, В. Ю. Михальченко. М., 2000.

Письменные языки мира: Российская Федерация. Социолингвистическая энциклопедия. Кн. 2. / отв. ред. В. М. Солнцев, В. Ю. Михальченко. М., 2003.

Сравнительно-историческая грамматика тюркских языков. Региональные реконструкции / отв. ред. Э. Р. Тенишев. М.: Наука, 2002.

Язык и общество. Энциклопедия / гл. ред. В. Ю. Михальченко. М.: «Азбуковник», 2016.

Языки народов России. Красная книга. Энциклопедический словарь-справочник / гл. ред. В. П. Нерознак. М.: «Academia», 2002.

Языки народов СССР. Т. 1. Индоевропейские языки / гл. ред. В. В. Виноградов. М.: Издательство «Наука», 1966.

Языки народов СССР. Т. 2. Тюркские языки / гл. ред. В. В. Виноградов. М.: Издательство «Наука», 1966.

Языки народов СССР. Т. 3. Финно-угорские и самодийские языки / гл. ред. В. В. Виноградов. М.: Издательство «Наука», 1966.

Языки народов СССР. Т. 4. Иберийско-кавказские языки / гл. ред. В. В. Виноградов. М.: Издательство «Наука», 1967.

Языки народов СССР. Т. 5. Монгольские, тунгусо-маньчжурские и палеоазиатские языки / гл. ред. В. В. Виноградов. М.: Ленинградское отделение издательства «Наука», 1968.

Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. Т. I. А-И. / ред. В. Н. Ярцева и др. М.: Издательство «Наука», 2001.

Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. Т. II. К-Р. / ред. В. А. Виноградов и др. М.: Издательство «Наука», 2001.

Языки Российской Федерации и соседних государств. Энциклопедия в 3-х томах. Т. III. С-Я. / ред. В. А. Виноградов и др. М.: Издательство «Наука», 2005.

Adler A. Language, or Dialect, That Is the Question. How Attitudes Affect Language Statistics Using the Example of Low German // *Languages*. 2021. 6:40.

Agard F. B. Language and dialect: some tentative postulates // *Linguistics*. 1971. 65. P. 5-24.

Ali M. H. The identity controversy of religious minorities in Iraq: the crystallization of the Yazidi identity after 2003 // *British Journal of Middle Eastern Studies*. 2020. 47:5. P. 811-831.

Blair F. Survey on a shoestring: A manual for small-scale language surveys. Dallas: Summer Institute of Linguistics and University of Texas at Arlington, 1990.

Bourgeois-Gironde S., Ginsburgh V., Hassani H., Weber S. A Lingua Franca for Kurdish Population. London: Centre for Economic Policy Research, 2021.

Campbell L., Belew A. (eds.). *Cataloguing the World's Endangered Languages*. London, New York: Routledge, 2018.

Casad E. H. Dialect intelligibility testing. Summer Institute of Linguistics Publications in Linguistics and Related Fields, 38. Norman: Summer Institute of Linguistics of the University of Oklahoma, 1974.

Catalogue of Endangered Languages: сайт. 2021. University of Hawaii at Manoa. URL: <http://www.endangeredlanguages.com> (дата обращения 20.11.2021).

- Chalvin A. La construction de l'identité des Setos (1920–1940) // *Études finnoougriennes*. 2011. T. 43. P. 119-144.
- Comrie B. *The languages of the Soviet Union*. Cambridge: Cambridge University Press, 1981.
- Dalby D. *Linguasphere register of the world's languages and speech communities*. Hebron: Linguasphere Observatory, 2000.
- Delsing L.-O., Åkesson K. L. *Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska*. Copenhagen: Nordiska ministerrådet, 2005.
- Doerfer G. Ist Kur-Urmisch ein nanaischer Dialekt? // *Ural-Altäische Jahrbücher*. 1975. 47. P. 51-63.
- Drettas G. *Aspects pontiques*. Paris: Association de Recherches Pluridisciplinaires, 1997.
- Dyen I. *A lexicostatistical classification of the Austronesian languages*. Baltimore: Waverly Press, 1965.
- Eberhard D. M., Simons G. F., Fennig C. D. (eds.). *Ethnologue: Languages of the World*. Twenty-fourth edition. Dallas, Texas: SIL International, 2021. URL: <http://www.ethnologue.com> (дата обращения 20.11.2021).
- Fortescue M. *Comparative Nivkh Dictionary*. Munchen: Lincom, 2016.
- Goltz R., Kleene A. *Niederdeutsch // Handbuch der Sprachminderheiten in Deutschland* / eds. R. Beyer, A. Plewnia. Tübingen: Narr, 2020. P. 171-226.
- Gooskens C. The contribution of linguistic factors to the intelligibility of closely related languages // *Journal of Multilingual and Multicultural Development*. 2007. 28:6. P. 445-467.
- Gooskens C. *Experimental Methods for Measuring Intelligibility of Closely Related Language Varieties* // *Handbook of Sociolinguistics* / eds. R. Bayley, C. Lukas. Oxford: Oxford University Press, 2013. P. 195-213.
- Gooskens C., Heeringa W. 2004. Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data // *Language Variation and Change*. 2004. 16. P. 189-207.
- Gooskens C., Beijering K., Heeringa W. Phonetic and lexical predictors of intelligibility // *International Journal of Humanities and Arts Computing*. 2008. 2 (1–2). P. 63-81.
- Greenberg J. H. (ed.). *Universals of languages*. Cambridge, Mass.: MIT Press, 1963.
- Gruzdeva E., Janhunen J. Revitalization of Nivkh on Sakhalin // *The Routledge Handbook of Language Revitalization* / eds. L. Hinton, L. Huss, G. Riche. New York: Routledge, 2018. P. 464-472.
- Gruzdeva E. On the diversification of Nivkh varieties // *The 4th Annual Meeting of Japan Association of Northern Language Studies, Hokkaido Museum of Northern Peoples, 6-7 November, Abashiri, Japan*. 2021.
- Hammarström H. Counting Languages in Dialect Continua Using the Criterion of Mutual Intelligibility // *Journal of Quantitative Linguistics*. 2008. 15 (1). P. 34-45.

Hammarström H., O'Connor L. Dependency-sensitive typological distance // Approaches to measuring linguistic differences / eds. L. Borin, A. Saxena. Berlin: De Gruyter Mouton, 2013. P. 329-352.

Hammarström H., Forkel R., Haspelmath M., Bank S. Glottolog 4.4. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2021. URL: <http://glottolog.org> (дата обращения 20.11.2021).

Harrison D. K. Language endangerment among the Tofa // Cultural Survival Quarterly. April 2003. P. 53-55.

Harrison D. K., Anderson G. D. S. Tofa language change and terminal generation speakers // Typological Studies in Language: Lessons from documented endangered languages / eds. D. K. Harrison, D. S. Rood, A. Dwyer. Amsterdam, Philadelphia: John Benjamins Publishing Co., 2008. P. 243-270.

Hassanpour A. Nationalism and language in Kurdistan, 1918-1985. San Francisco: Mellen Research University Press, 1992.

Haugen E. Dialect, Language, Nation // American Anthropologist New Series. 1966. 68:4. P. 922-935.

Haugen E. Semicommunication: The language gap in Scandinavia // Sociological Inquiry. 1966. 36(2). P. 280-297.

Homola P., Kuboň V. A structural similarity measure // Linguistic distances / eds. J. Nerbonne, E. Hinrichs. Sydney: ACL, 2006. P. 91-99.

Holman E. W., Wichmann S., Brown C. H., Velupillai V., Müller A., Brown P., Bakker D. Explorations in automated language comparison // Folia Linguistica. 2008. 42. P. 331-354.

James C. J., Cheesman M. F., Cornelisse L., Miller L. T. Response times to sentence verification tasks (SVTS) as a measure of effort in speech perception // Fifth Australian International Conference on Speech Science & Technology. 1994. P. 600-605.

Janhunen J. Tungusic: an endangered language family in Northeast Asia // International Journal of the Sociology of Language. 2005. 173. P. 37-54.

Kamusella T. The History of the Normative Opposition of "Language versus Dialect": From Its Graeco-Latin Origin to Central Europe's Ethnolinguistic Nation-States // Colloquia Humanistica. 2016. 5. P. 164-188.

Kaufman T. The native languages of Latin America: General remarks // Atlas of the world's languages / eds. C. Moseley, R. E. Asher R. E. London: Routledge, 1994. P. 31-33.

Kloss H. Abstand languages and Ausbau languages // Anthropological Linguistics. 1967. 9. P. 29-41.

Kluge A. RTT retelling method: An alternative approach to interdialectal comprehension testing // Proceedings of methods XIII: Papers from the Thirteenth International Conference on Methods in Dialectology / eds. B. Heselwood, C. Upton. Frankfurt: Bamberg, 2008. P. 174-183.

Koryakov Y. Atlas of the Caucasian languages. Moscow, 2002.

Koryakov Y. 2013. Convergence and divergence in the classification of Dargwa languages. In: 46th Annual Meeting of the Societas Linguistica Europaea (SLE 2013). 18–21 September 2013. Book of abstracts. Part 1. Split: University of Split.

Koryakov Y., Verhees S. Linguistic geography of East Caucasian languages. Lecture, 03.11.2021 // The online course on East Caucasian (alias Nakh-Daghestanian) languages.

Lander Y. Relativization in Shapsug Adyghe // Rice Working Paper in Linguistics. 2. P. 75-91.

Malyshev V., Malysheva V., Gutz A., Novaya I., Panina A., Yurkova A., Clifton J. M., Tiessen C. The Sociolinguistic Situation of the Dargwa in Dagestan // SIL Electronic Survey Reports 2019-011. Dallas, TX: SIL International, 2019.

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

Moseley C. (ed.). Encyclopedia of the World's Endangered Languages. London, New York: Routledge, 2007.

Moseley C. (ed.). Atlas of the World's Languages in Danger, 3rd edn. Paris: UNESCO Publishing, 2010. URL: www.unesco.org/languages-atlas/en/atlasmap.html (дата обращения: 20.11.2021).

Munkácsi B., Kálmán B. Wogulisches Wörterbuch. Budapest: Akadémiai Kiadó, 1986.

Nerbonne J., Wiersma X. A measure of aggregate syntactic distance // Linguistic distances / eds. J. Nerbonne, E. Hinrichs. Sydney: ACL, 2006. P. 82-90.

Pajusalu K. Estonian dialects // Estonian Language / ed. M. Ereht (ed.). Linguistica Uralica, supplementary series. Vol. 1. 2003. P. 231-272.

Polinsky M. Crimean Tatar and Krymchak: classification and description // The Non-Slavic languages of the USSR: linguistic studies / ed. H. I. Aronson. Chicago Linguistic Society, University of Chicago, 1992. P. 157-188.

Polivanov K., Kazakevich O., Serdobolskaya N., Khalilova Z., Budyanskaya E., Evstigneeva A., Mishchenkova K., Mordashova D., Pokrovskaya S., Renkovskaya E. The Website «Minority Languages of Russia»: Visualizing Language Data // Proceedings of the Linguistic Forum 2020: Language and Artificial Intelligence. CEUR Workshop Proceedings. 2852.

Rantala L., Sergina A. Áhkkila sápmelaččat: oanehis muitalus sámejoavkku birra, man mañimuš sámeielalaš olmmoš jámii 29.12.2003, s. 79 // Lapin yliopiston kasvatustieteellisiä raportteja. 8. Rovaniemi: Lapin yliopisto, 2009.

Salminen T. A key to Enets morphology. (Morphologisches Wörterbuch des Enzischen von Tibor Mikola) // Finnisch-Ugrische Forschungen. 1997. 54. P. 213-219.

Sammalahti P. The Saami languages: An Introduction. Karasjok: Davvi Girji, 1998.

Schüppert A. Origin of Asymmetry: Mutual Intelligibility of Spoken Danish and Swedish. PhD dissertation, University of Groningen, 2011.

Sejane E., Eger S. Semantic typologies by means of network analysis of bilingual dictionaries // Approaches to measuring linguistic differences / eds. L. Borin, A. Saxena. Berlin: De Gruyter Mouton, 2013. P. 447-474.

Siegl F., Rießler M. Uneven Steps to Literacy // Cultural and Linguistic Minorities in the Russian Federation and the European Union. Multilingual Education / eds. H. Marten, M. Rießler, J. Saarikivi, R. Toivanen. Vol 13. Cham, 2015. P. 189-230.

Singh A. K., Surana H. Can corpus based measures be used for comparative study of languages? // Proceedings of Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology. P. 40-47.

Starr A., Stadler J. Principles of sociolinguistic surveys: Actualities of Cameroon. Yaoundé: Summer Institute of Linguistics, 1990.

Tang C., Van Heuven V. J. Mutual intelligibility and similarity of Chinese dialects: Predicting judgments from objective measures // Linguistics in the Netherlands 24 / eds. B. Los, M. Van Koppen. Amsterdam: John Benjamins, 2007. 223-234.

van Hout R., Münsterman H. Linguistische afstand, dialect en attitude // Gramma. 1981. 5. 101-123.

Van Rooy R. Language or Dialect? The History of a Conceptual Pair. Oxford: Oxford University Press, 2020.

Wang W. S. Y. Glottochronology, Lexicostatistics, and Other Numerical Methods // The Encyclopedia of Language and Linguistics, Vol. 3 / ed. R. E. Asher. Oxford: Pergamon Press, 1994. P. 1445-1450.

Wälchli B., von Wandelfels R. Measuring morphosemantic language distance in parallel texts // Approaches to measuring linguistic differences / eds. L. Borin, A. Saxena. Berlin: De Gruyter Mouton, 2013. P. 475-506.

Wichmann S., Holman E. W., Brown C. H. (eds). The ASJP Database (version 18). 2018. URL: <http://asjp.clld.org/> (дата обращения 20.11.2021).

Wichmann S. How to Distinguish Languages and Dialects // Computational Linguistics. 2020. 45 (4). P. 823-831.

Wurm S. A. (ed.). Atlas of the World's Languages in Danger of Disappearing. 1st edition. UNESCO, Paris, 1996.

Wurm S. A. (ed.). Atlas of the World's Languages in Danger of Disappearing. 2nd edition. UNESCO, Paris, 2001.

Zhivlov M. Annotated Swadesh wordlists for the Ob-Ugric group (Uralic family). The Global Lexicostatistical Database. 2014. URL: <https://starlingdb.org/cgi-bin/response.cgi?root=new100&morpho=0&basename=new100\ura\oug&limit=-1> (дата обращения 20.11.2021).

11. Приложения

- [Приложение 1. “Предварительный список языков России”](#)

- [Приложение 2. “Идиомы, проблемные с точки зрения отнесения к языку или диалекту”](#)
- [Приложение 3. “Экстратерриториальные языки”](#)
- [Приложение 4. “Заснувшие языки”](#)