

РУССКИЙ ЯЗЫК В ДАГЕСТАНЕ: ПРОБЛЕМЫ ЯЗЫКОВОЙ ИНТЕРФЕРЕНЦИИ¹

Даниэль М. А. (misha.daniel@gmail.com),
Добрушина Н. Р. (nina.dobrushina@gmail.com)

НИУ ВШЭ, Москва, Россия

Дагестан представляет собой территорию, где стремительно растет число носителей русского языка. В последние десятилетия русский язык занял здесь место, которое до того времени было вакантно — он стал лингва франка всего Дагестана. При этом, поскольку русский язык усваивается местными жителями в интенсивном контакте с местными языками и одновременно все больше и больше, в своем локальном варианте, приобретает (само)идентификационные функции, он обладает многими характерными особенностями на уровне фонетики, морфологии, синтаксиса и лексики. Цель настоящей статьи — обсудив социолингвистические и языковые особенности сельского варианта русского языка в Дагестане, предположительно разделить их на те, которые мотивированы интерференцией, и те, которые мотивированы феноменом *imperfect learning*, а также, в заключении, обосновать необходимость создания корпуса дагестанского русского как исследовательского инструмента.

Ключевые слова: языковая интерференция, русский язык, языки Дагестана, региональные языковые варианты, разработка корпусных ресурсов

A CORPUS OF RUSSIAN AS L2: THE CASE OF DAGHESTAN²

Daniel M. A. (misha.daniel@gmail.com),
Dobrushina N. R. (nina.dobrushina@gmail.com)

NRU HSE / MSU, Moscow, Russia

¹ Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ в 2013 году.

² This article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE).

In Daghestan, the number of Russian speakers has been dramatically increasing over the last few decades. Russian has assumed the functional niche previously vacant in this extremely multilingual setting, becoming the first ever lingua franca of the region as a whole. Russian is acquired in a situation of strong interaction with local languages and shows contact properties on various linguistic levels: phonetics, morphology, syntax and lexicon. Its regional variant is also visibly developing as a self-identification device. The aim of this paper to discuss some (socio)linguistic properties of this idiom, attribute them either to interference or to imperfect learning, and to argue for building a corpus of Daghestanian Russian.

Keywords: language interference, imperfect learning, Russian as L2, East Caucasian, regional variants, corpus development

Роль и усвоение русского языка в Дагестане

В Дагестане более сорока языков (точное число зависит от выделения различных языков в диалектно-языковых кластерах, в первую очередь в даргинском кластере), и дистанция между этими языками очень велика. Хотя языки семьи обладают чертами нетривиального структурного сходства, даже те из них, которые являются близкородственными, такие как лезгинский и табасаранский, практически полностью исключают возможность взаимопонимания. Глубина родства языков этой семьи сопоставима с глубиной родства индоевропейских языков.

Есть данные о том, что в 19 веке в разных регионах Дагестана в качестве лингва франка функционировали разные языки. Роль языков межэтнической коммуникации часто выполняли тюркские языки — ногайский, кумыкский, азербайджанский, носители которых обитали на нижних территориях Дагестана (эти языки, например, названы в Chirikba 2008: 74). Исследователями описана модель вертикального билингвизма, при которой языки распространяются снизу вверх: жители верхних, высокогорных сел владеют языками нижних, а нижние языками верхних не владеют (Wixman 1980, Волкова 1974, Nichols forthcoming). У этого явления есть много причин, в их числе основной вектор передвижений (жители верхних сел спускались по разным хозяйственным нуждам вниз, но не наоборот), более высокий социальный статус земледельцев по сравнению со скотоводами (Карпов, Капустина 2012; в верхнем Дагестане занимались почти исключительно овцеводством), взаимная близость этих тюркских языков, которая позволяет тому, кто знает кумыкский, понимать азербайджанский и ногайский, и наоборот.

В некоторых регионах роль лингва франка выполнял аварский язык. Кроме того, на территории Дагестана функционировал арабский язык, который занимал место письменного литературного языка. По мнению Вихмана (Wixman 1980: 115), в 19 веке в Дагестане преподавание в школах происходило только на арабском, он же был письменным языком гражданских служб. Арабский был языком интеллигенции вплоть до революции, и первые лидеры Коммунистической партии Дагестана, принадлежавшие к числу наиболее образованных местных людей, тоже владели арабским.

Как бы то ни было, до середины 20 века в Дагестане не было языка, который объединял бы всю его территорию. Это исключительное место сумел занять русский язык. Процесс русификации Дагестана, по-видимому, начинается с 19 века, когда в 1820 году в результате военных действий Дагестан был присоединен к России. Есть данные, что в 1897 году на территории Дагестана было 6,3% русских, а к 1913 году их численность увеличилась в 2,2 раза (Ибрагимов 1978, цит. по Магомедханов 2008). Каковы каналы усвоения русского языка в сегодняшнем Дагестане? Ситуация в горных селах и в городах принципиально различна.

В горах Дагестана этнических русских практически нет. В высокогорных селах, в которых нам пришлось побывать в последние годы, мы встретили русскую жительницу лишь однажды: это была очень пожилая учительница, которая в 50-е годы приехала по распределению в Дагестан работать в школе, вышла замуж и осталась жить в селе (Чародинский район). Присутствие в сельских школах русских учителей было более характерно для 50–70-х годов (см. об этом Добрушина 2008), но часто речь шла об одном учителе на селение.

По данным опроса, произведенного в селениях Арчиб, Читтаб и Шалиб в 2008–2010 годах, среди людей, родившихся до 1919 года, некоторая степень владения русским языком была характерна для 18% населения Арчиба, 32% населения Читтаба и 40% населения Шалиба. Важно отметить, что у нас нет никаких инструментов для того, чтобы определить уровень владения: вполне вероятно, что во многих случаях речь идет о способности объясниться на бытовом уровне. К тому же эти данные получены от родственников, которые помнили своих бабушек и дедушек уже в преклонном возрасте, то есть факт владения русским языком теми, кто родился после 1919 года, относится к периоду 60–70 годов. На протяжении всего 20-го века знание русского языка постоянно росло, так что в настоящее время по-русски говорят практически все, за исключением инвалидов, никогда не выезжавших из села, и очень пожилых людей. На рисунке 1 представлена динамика владения русским языком в высокогорном селении Арчиб.

В 30–50 годы основным источником знания русского языка была школа, общение с русскими учителями и выезды из села на русскоговорящие территории. Сегодня важным каналом является телевидение, которое присутствует везде с 80–90-х годов, причем популярными каналами являются общероссийские. Практика в разговорном русском возникает в селе летом, когда на каникулы приезжают дети, которые живут в городах. Полевые исследования говорят о том, что во многих высокогорных селах дети совсем не говорят по-русски до школы, а полноценное, свободное владение приходит уже после школы, когда молодежь начинает получать образование вне села, юноши попадают в армию или начинают выезжать на заработки.

Как мы видим, единственным источником русского языка в том варианте, в каком он используется в тех регионах, где русский язык является первым или единственным, оказывается телевизор. Все остальные источники русского языка связаны с местными носителями, для которых русский язык часто представляет собой L2 (второй язык).

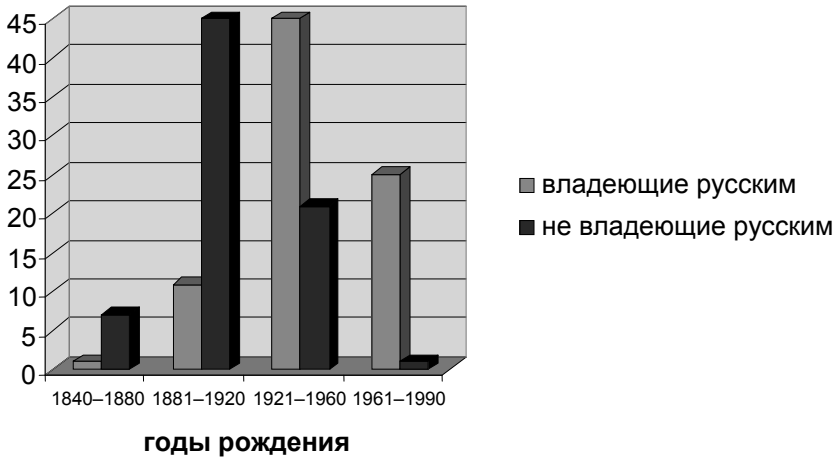


Рис. 1. Владение русским языком в дагестанском селении Арчиб по результатам опроса

В городах и, вероятно, в низинных селах (о них у нас меньше информации) ситуация иная. Русский язык получил исключительное распространение в городской жизни, поскольку города населены представителями всех дагестанских этносов, и русский функционирует как язык межэтнического общения. Дети, которые сегодня рождаются в городах, как правило, не владеют никакими этническими языками. Достаточно распространена также ситуация ограниченного владения этническим языком как L2, когда, например, ребенок общается по-аварски или по-лакски с бабушкой или по крайней мере понимает обращенную к нему речь; слышит его от односельчан, когда приезжает в село на каникулы, но сам говорить на языке не может.

Каковы каналы усвоения русского языка в дагестанских городах? Несмотря на его распространенность, главная особенность сохраняется: в окружении городского ребенка практически нет источников русского языка в его исходном варианте. Городской ребенок усваивает русский язык от родителей, которые зачастую переехали в города из сел и являются носителями русского как L2 (процесс миграции сверху вниз сегодня исключительно активен). Наблюдения показывают, что такие родители предпочитают говорить с ребенком на русском, считая этот язык более полезным для будущей жизни. Родители могут быть и городскими жителями по рождению, и тогда русский язык является для них L1. Русский язык является языком общения в детском саду и в школе и предметом школьного обучения. При этом этнических русских в городах Дагестана немного, по официальной статистике — 4,7% в 2002, 3,6% в 2010.

Таким образом, в городах единственным источником русского языка в том варианте, в каком он используется на русскоговорящих территориях России, вновь оказывается телевизор. Разница с селами состоит в том, что в окружении ребенка значительно больше таких людей, для которых русский является родным языком, и в том, что вокруг себя он слышит преимущественно русскую речь.

Итак, специфика сегодняшнего Дагестана состоит в том, что он является территорией стремительно растущего русскоязычного населения при почти полном отсутствии этнических русских. Неудивительно, что в этой ситуации формируется специфический вариант русского языка, обладающий определенными отличительными чертами. Можно было бы обозначить этот вариант термином *этнолект*. Согласно одному из наиболее популярных определений этнолекта, это “varieties of a language that mark speakers as members of ethnic groups who originally used another language or distinctive variety” (Clyne 2000:291). Применение данного определения осложнено тем фактом, что русский язык Дагестана формируется на почве многих различных этнических групп. Но поскольку механизмы, вероятно, отчасти схожи с теми, о которых пишет Клайн, мы условно примем этот термин для нашей ситуации.

Важным компонентом определения этнолекта является то, что он служит инструментом для создания групповой идентичности. Мы не можем однозначно ответить на вопрос о том, служит ли дагестанский русский для целей идентификации, поскольку не проводили никаких систематических исследований такого рода, хотя отдельные указания на это имеются (юмористические передачи дагестанцев и о дагестанцах, обсуждение слов «дагестанского языка» на форумах, составляемые дагестанцами «словарики» локального варианта и проч.). Выскажем некоторые предположения о механизмах этого процесса.

Стремительная урбанизация Дагестана и утрата знания местных языков создает препятствия для конструирования этнической идентификации у молодых городских дагестанцев: они лишены этнического языка как основной опоры этничности. В этой ситуации русский вариант Дагестана оказывается полезным субститутом. Ему соответствует этничность, которая на данный момент служит маркером «внешней» идентификации: вне Дагестана его жители называют себя дагестанцами (в то время как внутри Дагестана имеют хождение такие маркеры, как аварцы, даргинцы, лезгины и так далее). Такая этничность не имеет и не может иметь собственного языка, который отличал бы ее от других групп. Представляется, что это место может занять региональный вариант русского.

Как уже было сказано, трудность с определением дагестанского русского языка как этнолекта связана с тем, что он существует во многих вариантах. В самом деле, русский вариант Дагестана развивается в тесном контакте с местными языками. Этих языков около сорока. Значит, русская речь носителей каждого из этих языков, особенно в селах, может иметь какие-то особые черты, обусловленные именно этим языком. Это предположение необязательно верно: языки нахско-дагестанской семьи обладают значительным структурным сходством, так что не исключено и то, что больших расхождений между разными вариантами русского нет. К тому же тот русский язык, который функционирует в городе, вполне вероятно, представляет собой наддиалектный вариант — результат выравнивания сельских вариантов. Тем не менее исключить такую внутреннюю неоднородность регионального русского невозможно.

Для того чтобы перевести эти предположения в разряд утверждений, необходимы обширные исследования дагестанского русского языка в его сельских и городских вариантах. В этой статье, также как и в нескольких предшествующих

публикациях (Даниэль, Добрушина 2010; Daniel, Dobrushina, Kniazev 2011), мы сосредоточились на первой задаче. В отличие от предыдущих публикаций, здесь мы не только представляем типы нестандартных грамматических явлений в собранных нами текстах, но и предпринимает попытку их классификации как вызванных одной из двух возможных причин: интерференции с первым языком или так называемого «неполного выучивания» (imperfect learning).

Грамматические особенности русского языка сельчан

Языковой материал, обсуждаемый в настоящей статье, опирается на доступные нам данные дагестанского русского языка в трех селениях: Арчиб, Читтаб и Шалиб. Эти три селения находятся на расстоянии пешего хода (примерно полтора часа от одного до другого), а их жители активно контактируют друг с другом. В каждом из селений свой родной язык (арчинский, аварский и лакский соответственно; все три языка принадлежат к разным ветвям нахско-дагестанской семьи). Традиционно для этих сел был характерен некоторый уровень знания языка соседей. В настоящее время есть тенденция к тому, что языком коммуникации с соседями становится русский, во всяком случае для молодежи.

Лексические заимствования, во многих отношениях представляющие собой интерес, ниже не рассматриваются; разбор нескольких ярких случаев (см. в Daniel, Dobrushina, Kniazev 2011). Там же дается более подробный обзор типов замеченных нами грамматических и синтаксических особенностей. В настоящей краткой заметке мы хотели бы акцентировать внимание на том, что наблюдаемые особенности могут проистекать из двух причин, или, точнее, иметь две составляющие: интерференция русского языка с первым языком говорящего либо эффект «неполного выучивания» (imperfect learning) второго языка. Эти составляющие могут комбинироваться, но часто они относительно четко различимы. В случае интерференции речь идет о проекции грамматических особенностей родного языка говорящего на структуры русского языка как L2. В случае «неполного выучивания» языковая система L2 оказывается «упрощена» вне зависимости от грамматических особенностей первого языка говорящего. Наконец, третья ситуация в каком-то смысле промежуточна — imperfect learning поддерживается отсутствием в родном языке того или иного грамматического явления, характерного для русского языка.

Идея неполного выучивания отнюдь не предполагает, что говорящий испытывает трудности при использовании второго языка. Еще раз подчеркнем, что подавляющее большинство живущих в селе (не говоря о городе) дагестанцев, с которыми мы встречаемся, говорят по-русски так же свободно, как и любой носитель русского языка как первого (и единственного). «Ошибки» у говорящих на таком варианте языка обычно интерпретируются как указание на то, что литературный язык просто не является для них целевой системой (target language) в процессе овладения языком. Некоторые исследователи полагают, что феномен неполного выучивания занимает центральное место в истории многих языков.

Сразу же можно отметить, что грамматические и иные особенности, наблюдаемые в наших текстах, конечно, никогда не являются абсолютно последовательными. Нестабилизированные языковые варианты (подобно диалектам, существующим в контакте с литературным языком, или креольским языкам, погруженным в посткреольский континуум) характеризуются более или менее заметными колебаниями между двумя языковыми системами — стандартным вариантом носителей языка как L1 и структурными элементами первого языка носителей этого же языка как L2. Примеры таких колебаний иногда приводятся ниже вместе с примерами нестандартных структур.

Итак, мы предлагаем различать интерференцию как «положительную» проекцию элементов языковой системы L1 на L1 (interference); неполное выучивание как элиминирование элементов языковой системы L2, отсутствующих в L1 (imperfect learning ~ interference); и неполное выучивание как результат выравнивания языковых особенностей L2 вне зависимости от структуры L1 (imperfect learning). Приведем наиболее яркие примеры для каждого из типов.

Interference: эвиденциальное *оказывается*

Ярким примером интерференции является чрезвычайно характерное для многих носителей употребление формы *оказывается*:

- (1) Да. Этот самый — он умер покойный — участник войны был Мусаев Магомед. Хороший, грамотный. Хорошо он знал русский язык. Он **оказывается** в детстве ходил туда в лакское селение, дома строить, помощником. Каменщика помощником. А оттудова он убежал, попал среди русских, и он говорил, что у него два сына в Волгограде. И связи не было. Такие тоже случались. Дальше... наши люди знали и русский язык. Вот наша соседка рассказывала, она младше матери была, **оказывается**, во время гражданской войны, восемнадцатом-двадцатых годах пришли солдаты до Кубатль. И там они начали сыпать зерно лошадям. А эта женщина, **оказывается**, знала русский язык. И начала она стыдить. Люди здесь голодают, а вы зерно лошадям даете. Она поругала их, и они говорят перестали зерно брать. Это уже женщина, не мужчина. Я думаю, их наверное послали в Сибирь или что-то наверно такое было. Тоже, значит, русский язык знали люди. А азербайджанский почти второй человек знал у нас.
- (2) Кочубей это черный рынок называли раньше. **Оказывается** на Кавказе был рынок в этом месте. Там продавали девушек. и поэтому говорят называли черным рынком его. такое место **оказывается** было.
- (3) [А сюда кто-нибудь приходил учиться?] Приходили. Вот к этому Мамма дибиру **оказывается** приходили. Даже он предчувствовал, что умрет. У него детей не было. Он сказал хозяйке — я завтра утром умру. А похоронить

приедут люди из Советского района. Шамилевский район. В самом деле. Утром **оказывается** он умер. Вышла эта хозяйка сказать — а оттудова пришли вот эти аварцы.

Очевидна своеобразность такого употребления. В литературном русском языке это вводное слово используется для введения информации, которую говорящий представляет как неожиданную для себя (так называемый *миратив*). В приведенном же отрывке рассказа это значение очевидно отсутствует, и для носителя литературного русского языка такие употребления оказываются немотивированными.

Такое употребление является результатом интерференции с L1. В нахско-дагестанских языках вообще и в арчинском, аварском и лакском языках, в частности, присутствует категория эвиденциальности (иначе косвенной засвидетельствованности или заглазости). Эвиденциальный показатель (или конструкция) указывают, что говорящий сам не является свидетелем передаваемых им событий. По крайней мере в арчинском языке эта категория является строго обязательной (Кибрик 1977). Говоря по-русски, наши собеседники испытывают давление L1 и подбирают средство выражения в качестве функционального заместителя эвиденциальности. Интересно отметить, что выбор их, по-видимому, не случаен: категория миратива, на первый взгляд логически независимая от эвиденциальности, обнаруживает типологическую связь с последней (Плунгян 2011).

Interference: опущение предлогов

Несколько менее очевидный случай, но не менее яркая черта речи наших собеседников — это отсутствие предлогов.

- (4) Курбамагомед он же живёт/ он **Качалибе** живёт
(Качалиб — название населенного пункта)
- (5) Там много топономии вот/ топономики/ эт самое/ названий мест/ упомянуто вот **этом письме завещании**.
- (6) И// до этого конечно/ до этого надо идти **нашему устару**// Наш устар/ вот шейх Мамадибир/ ээ у него надо как бы разрешение брать// Ну// Наставник/ учитель как бы религиозный/ **вот/ я ему тоже пошёл**// в этом году/ двадцать четвёртого// марта/ этот день я помню. [устар — мастер, наставник]

Мы предполагаем, что эта черта является проявлением также языковой интерференции. Для языков Дагестана предлоги вообще не характерны (они используют послелогии); кроме того, очень большой инвентарь пространственных значений выражается собственно именным склонением. В зависимости от конкретного языка, более или менее многочисленные (иногда до нескольких

десятков) комбинации пространственных значений, такие как ‘на верхнюю поверхность X-а’ или ‘домой к X-у’, могут выражаться морфологическими средствами. Говоря по-русски, под давлением структуры L1 наши собеседники иногда опускают чуждое для нее аналитическое препозитивное служебное слово. Такое опущение не является последовательным; так, в том же интервью наш собеседник употребляет предлог в очень похожем контексте.

(7) ... брать разрешение/ что я вот так хочу/ с вашего разрешения/ вот так хочу/ пойти **к другому устару**/ который щас живёт/ жив/ вот// Вот **к нему** надо идти/ сделать товбу

Конечно, и в данном, и в других подобных случаях нельзя исключать наличие систематических собственно лингвистических факторов, влияющих на использование или опущение предлога. Судить об этом определенно однако нельзя по причине недостаточности доступного языкового материала.

Imperfect learning ← Interference: колебания возвратности

В текстах встречается большое число примеров с употреблением возвратной формы глагола в невозвратном контексте и наоборот; некоторые глагольные формы в литературном языке просто отсутствуют.

(8) Необразованные наши вот арчинки вот **попадают**ся в больницу. (A)

(9) Они тоже **гостились** у нас Цурибе. (Ч)

(10) День рождения **справились**, там документы были на окошке. (Ч)

(11) Если я здесь **родила**, отец-мать арчинцы, представляешь, я арчибском языке पहले говорила — как же я аварка, скажи? (A)

Здесь мы имеем дело с неполным выучиванием, мотивированным несоответствием между структурами L1 и L2. В русском языке единственный высокопродуктивный морфологический процесс, связанный с изменением аргументной структуры глагола — это понижение переходности, т. е. образование медиопассивных (возвратных) форм. Нахско-дагестанские языки, напротив, принадлежат к транзитивизирующим языкам, то есть языкам, явно предпочитающих повышение переходности — каузативизацию. Понижающие актантные деривации в этих языках либо отсутствуют, либо занимают сугубо периферийную позицию. (О делении языков на повышающие и понижающие переходности см. Nichols et al. 2004.) Таким образом, L1 не представляет никаких функциональных аналогов для данного явления L2, на которые говорящий мог бы опереться при порождении формы, и в использовании возвратных форм наблюдаются сильные колебания.

Imperfect learning: *именительный вместо родительного*

Для русского языка характерно употребление родительного падежа вместо именительного и винительного падежей в отрицательных контекстах. В первом из приведенных примерах такая замена в норме обязательна, но отсутствует в записанном примере. Во втором случае она в норме, наоборот, невозможна, но говорящий, по-видимому, интерпретирует отрицательную семантику контекста ('плохо знал' аналогично 'не знал') как триггер такой замены:

- (12) Азербайджан ехали люди/ которые// было тяжёлое положение// Здесь у них **поля' не было**
- (13) Вот// и он вот так начал// он плохо **знал аварского языка**// Ну/ русского вообще не знал//

Очевидно, что в этом случае говорить об интерференции с L1 не приходится: в дагестанских языках просто нет структур, которые могли бы следовать или не следовать этим правилам. Речь идет, так сказать, о неприменении (или применении в другом классе контекстов) правила, специфичного для грамматической системы L2.

Imperfect learning: *видовые колебания*

В некоторых случаях наблюдаются колебания в приписывании граммы вида. Для следующего контекста вполне вероятно, что мы имеем дело просто с видовой ошибкой:

- (14) А после восемьдесят пятого года/ когда уже религию как бы// не **запретили**/ вот// Щас все/ я же говорю/ все как бы начали/ и молиться/ и соблюдать пост// Особенно молодёжь

Однако в нескольких других примерах, обсуждаемых в (Daniel, Dobrushina, Kniazev 2011), речь идет об использовании совершенного вида для описания предельных ситуаций в узуальных контекстах, где в русском языке обычно происходит замена характерной для предельных ситуаций граммы совершенного вида на грамму несовершенного вида в значении хабитуалиса; возможно, что и в предыдущем примере использование совершенного вида связано с предельностью глагола. Снова отметим характерную для L2 непоследовательность: глагол *снимать* является, очевидно, предельным, но употреблен в форме узуального несовершенного вида.

- (15) Я в начальных классах училась, чуть старшие, кто седьмой, шестой класс учится, вот они вот так вот чухту, каз ходили, в школе не разрешают, они на дороге снимали, **спрятали** там.

Несмотря на существенные отличия видовой системы нахско-дагестанских языков от славянского вида, эти языки никак нельзя считать «невидовыми» и интерпретировать соответствующие примеры аналогично колебаниям возвратности. Если принять гипотезу о связи таких контекстов с предельностью, придется признать, что видовые системы L1 и L2 существуют тут вполне независимо; расширение контекстов употребления совершенного вида мотивировано семантическими факторами (предельность глагола).

Эти и другие (см. Даниэль, Добрушина 2010; Daniel, Dobrushina, Kniazev 2011) свойства исследуемых нами вариантов русского языка можно предварительно охарактеризовать с точки зрения их отнесения к интерференции или неполному выучиванию:

| | Интерференция | Неполное выучивание |
|-------------------------------------|---------------|---------------------|
| <i>оказывается</i> | + | – |
| нативные модели полисемии | + | – |
| левое ветвление | + | – |
| опущение предлогов | + | ? |
| нестандартное управление | + | ? |
| Асс (1 скл.) → Nom | +? | –? |
| колебания возвратности | + | + |
| Gen → Nom | – | + |
| рассогласование по роду | – | + |
| <i>было</i> как несогласуемая форма | – | + |
| Ipfv → Pfv | – | + |

Перспективы исследования: конпус дагестанского русского языка

Как уже несколько раз говорилось, окончательное решение многих центральных вопросов невозможно до тех пор, пока не будет собрано достаточно большое количество данных. Расшифрованных на сегодняшний день текстов сугубо недостаточно. Их формат не позволяет осуществлять поиск таким образом, чтобы можно было интерпретировать результаты в количественном отношении, для чего необходимы данные о числе и типах не только контекстов, где эти явления встретились, но и контекстов, где они могли бы встретиться, но отсутствуют. Как нам кажется, нам уже известны основные типы явлений, характерные для русского как L2 в сельских районах Дагестана; но у нас твердого представления о степени их распространенности. В некоторых случаях они могут коррелировать с определенными лингвистическими параметрами контекста; и только доступность исчерпывающих подборок примеров позволит подтвердить или опровергнуть предположительное отнесение этих явлений к интерференции и/или к неполному выучиванию.

Решить часть этих проблем призван корпус русского языка в Дагестане. В настоящий момент наши тексты состоят почти исключительно из социолингвистических интервью, которые брались в ряде горных сел. Значительная часть интервью была расшифрована А. Е. Дьячковой в 2011 году в рамках Программы фундаментальных исследований РАН. В настоящий момент расшифровано 17 текстов, общим объемом 7 часов звучания (около 50 тысяч словоупотреблений). Как и приведенные выше примеры, расшифровка записей осуществляется в формате устного подкорпуса Национального корпуса русского языка (www.ruscorgo.ru); снятые в этой публикации ударения в текстах корпуса присутствуют. В течение 2013 года планируется расшифровка еще 5 часов, разметка и вывеска текстов в открытом доступе.

Разметка корпуса должна включать как социолингвистический (пол, возраст, место жительства и первый язык носителя, его лингвистическая биография), так и собственно лингвистический компонент. К последнему относится класс «ошибки» и указание дополнительных сведений, необходимых для удобного и информативного поиска по корпусу. Например, для ошибок согласования и примеров левого ветвления должны указываться характеристики вершины; при опущении предлога должен указываться этот предлог и проч.

Поскольку тексты были записаны в селах, все они представляют русский язык как L2. В таблице приведен список селений, в которых производилась запись (с указанием L1 для каждого селения).

Таблица 1. Список селений

| село | L1 |
|-----------|---------------|
| Арчиб | арчинский |
| Читтаб | аварский |
| Шалиб | лакский |
| Ицари | ицаринский |
| Меgeb | меgebский |
| Хив | табасаранский |
| Маллакент | даргинский |
| Янгикент | кумыкский |

Увеличение объема расшифрованного корпуса позволит выяснить, каков статус описанных явлений в речи отдельных носителей; какие из них являются относительно стабильными, а какие носят скорее характер окказионализмов. Выше мы видели, что один и тот же носитель чаще всего использует и нормативную, и альтернативную конструкции; вероятно, ни один носитель не бывает вполне последовательным в следовании локальному варианту, но приверженность норме разные носители демонстрируют в разной степени.

Не менее интересным является вопрос о социальных параметрах варьирования тех или иных явлений: зависит ли их присутствие в речи и частотность от возраста, образования, пола говорящего? Какие именно явления характерны для всех категорий говорящих, а какие имеют социальную обусловленность?

Важно выяснить, есть ли такие явления в русской речи дагестанцев, которые коррелируют с их родным языком, например, встречаются в речи аварцев, но нехарактерны для носителей даргинского.

Наконец, помимо текстов сельского русского, необходимы записи городской русской речи тех, для кого русский язык является L1. Сопоставление особенностей дагестанского русского как L2 с дагестанским русским как L1 может дать представление о механизмах формирования нового варианта русского языка.

Литература

1. Волкова, Н. Г. (1974). Этнический состав населения Северного Кавказа в XVII-начало XX века. Москва: Наука.
2. Даниэль М. А., Добрушина Н. Р. (2010). Новые русские. // *Вопросы русского языкознания*. Вып. XIII. Фонетика и грамматика: настоящее, прошедшее, будущее. Стр. 141–158.
3. Карпов, Ю. Ю. & Капустина, Е. Л. (2011). Горцы после гор. Миграционные процессы в Дагестане в XX — начале XXI века: их социальные и этнокультурные последствия и перспективы. Санкт-Петербург.
4. Кибрик А. Е. (1977). Опыт структурного описания арчинского языка. Т. 1–3. Москва: МГУ.
5. Магомедханов М. М. (2007). Дагестанцы: вехи этносоциальной истории 19–20 вв. Издательство ДНЦ РАН, Махачкала.
6. Плунгян В. А. (2011). Введение в грамматическую семантику: Грамматические значения и грамматические системы языков мира. М.: РГГУ.
7. Chirikba, Viacheslav A. (2008). The problem of the Caucasian Sprachbund. In Pieter Muysken (ed.) *From linguistic areas to areal linguistics*. John Benjamins Publishing Company. pp. 25–94.
8. Clyne, Michael. (2000). *Lingua Franca and ethnolects in Europe and beyond*. *Sociolinguistica*, 14, pp. 83–89
9. Nichols Johanna, David A. Peterson, Jonathan Barnes. (2004). Transitivity and detransitivizing languages. *Linguistic Typology*, vol. 8, no. 2, pp. 149–211.
10. Daniel, Michael, Nina Dobrushina and Sergei Kniazev. (2011). Highlander's Russian: Case Study in Bilingualism and Language Interference in Central Dagestan. In: *Instrumentarium of Linguistics: Sociolinguistic Approach to Non-Standard Russian*. *Slavica Helsingiensia*, 40, pp. 65–93.
11. Nichols, Johanna, forthcoming. The vertical archipelago: altitude, typology, and sociolinguistics in mountain languages. In Peter Auer, Martin Hilpert, Anja Stukenbrock and Benedikt Szmrecsanyi, eds., *Space in Language and Linguistics*. Berlin: Mouton de Gruyter.
12. Thomason, Sarah, Terrence Kaufman. (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, University of California Press.
13. Wixman, Ronald. (1980). *Language Aspects of Ethnic Patterns and Processes in the North Caucasus*. Chicago: University of Chicago.

References

1. *Chirikba, Viacheslav A.* (2008). The problem of the Caucasian Sprachbund. In Pieter Muysken (ed.) *From linguistic areas to areal linguistics*. John Benjamins Publishing Company. pp. 25–94.
2. *Clyne, Michael.* (2000). Lingua Franca and ethnolects in Europe and beyond. *Sociolinguistica*, 14, pp. 83–89
3. *Nichols, Johanna, David A. Peterson, Jonathan Barnes.* (2004). Transitivity and detransitivizing languages. *Linguistic Typology*, vol. 8, no. 2, pp. 149–211.
4. *Daniel, Michael, Nina Dobrushina and Sergei Kniazev.* (2011). Highlander's Russian: Case Study in Bilingualism and Language Interference in Central Daghestan. In: *Instrumentarium of Linguistics: Sociolinguistic Approach to Non-Standard Russian*. *Slavica Helsingiensia*, 40, pp. 65–93.
5. *Nichols, Johanna,* forthcoming. *The vertical archipelago: altitude, typology, and sociolinguistics in mountain languages*. In Peter Auer, Martin Hilpert, Anja Stukenbrock and Benedikt Szmrecsanyi, eds., *Space in Language and Linguistics*. Berlin: Mouton de Gruyter.
6. *Thomason, Sarah, Terrence Kaufman.* (1988). *Language contact, creolization, and genetic linguistics*. Berkeley, University of California Press.
7. *Wixman, Ronald.* (1980). *Language Aspects of Ethnic Patterns and Processes in the North Caucasus*. Chicago: University of Chicago.
8. *Volkova, Natalja G.* (1974). *Etnicheskij sostav naselenija Severnogo Kavkaza v XVIII — nachale XX veka*. Moskva: Nauka.
9. *Daniel', Mikhail, Nina Dobrushina.* (2010). *Novye russkie*. // *Voprosy russkogo jazykoznanija*. Vyp. XIII. Fonetika i grammatika: nastojashhee, proshedshee, budushhee. Str. 141–158.
10. *Karpov, Jurij Ju., Ekaterina L. Kapustina.* (2011). *Gorcy posle gor. Migracionnyje processy v Dagestane v XX—nachale XXI veka: ix sotsial'nye i etnokul'turnye posledstvija i perspektivy*. Sankt-Peterburg: Peterburgskoe Vostokovedenie.
11. *Kibrik, Aleksandr E.* (1977). *Opyt strukturnogo opisanija archinskogo jazyka*. T. I. Leksika. Fonetika. T. II. Taksonomicheskaja grammatika. Izdatel'stvo Moskovskogo universiteta.
12. *Magomedkhanov, Magomed.* (2007). *Dagestancy: vekhi etnosocial'noj istorii 19–20 vv.* Izdatel'stvo DNC RAN, Makhachkala.
13. *Plungjan V. A.* (2011). *Vvedenie v grammaticheskiju semantiku: Grammaticheskie znachenija i grammaticheskie sistemy jazykov mira*. M.: RGGU.