

УДК 81'27

## ЛИНГВИСТИЧЕСКИЙ КОРПУС «ТОМСКИЙ РЕГИОНАЛЬНЫЙ ТЕКСТ»: ЯВЛЕНИЯ ЯЗЫКОВОЙ ИНТЕРФЕРЕНЦИИ В РУССКОЙ РЕЧИ БИЛИНГВОВ ТОМСКА И ТОМСКОЙ ОБЛАСТИ

<sup>1</sup>Веснина Г.Ю., <sup>2</sup>Кустова Г.И.

<sup>1</sup>Томский государственный университет, Томск, e-mail: galina.y.vesnina@gmail.com;

<sup>2</sup>Институт русского языка им. В.В. Виноградова РАН, Москва, e-mail: galinak03@gmail.com

Статья посвящена обзору проекта подкорпуса устной и письменной русской речи билингвов, проживающих на территории Томского региона. Цель создания данного подкорпуса – отразить специфику русского языка в условиях его существования в билингвальных ситуациях. На материале русско-татарского билингвизма были выявлены интерферентные явления в морфологии при выборе родового окончания, образовании притяжательных форм прилагательных и местоимений, образовании возвратных глагольных форм. Синтаксическая интерференция проявляется в устойчивом порядке слов типа SOV. На лексическом уровне обнаружены заимствования из татарского языка в русский и нарушения лексической сочетаемости.

**Ключевые слова:** корпусная лингвистика, лингвистический корпус, региональная лингвистика, региолект, билингвизм, языковая интерференция, русский язык, татарский язык

## LINGUISTIC CORPUS «TOMSK REGIONAL TEXT»: LANGUAGE INTERFERENCE IN RUSSIAN SPEECH OF BILINGUALS FROM TOMSK AND TOMSK REGION

<sup>1</sup>Vesnina G.Y., <sup>2</sup>Kustova G.I.

<sup>1</sup>Tomsk State University, Tomsk, e-mail: galina.y.vesnina@gmail.com;

<sup>2</sup>The V.V. Vinogradov Russian Language Institute of the Russian Academy of Sciences, Moscow, e-mail: galinak03@gmail.com

The paper explores the subcorpus of spoken and written Russian language of bilingual Russian-speakers, which is a part of the corpus «Tomsk regional text». The purpose of creating a special subcorpus of Russian text and speech produced by bilinguals is to show specifics of Russian language use in bilingual situations. The material revealed evidence of the morphological interference in category of gender, possessive forms of adjectives and possessive pronouns, reflexive verb forms. Revealed syntactic interference shows in constant use of SOV word order. Lexical interference appears in loan words from Tatar into Russian and collocation errors.

**Keywords:** corpus linguistics, linguistic corpora, markup, regional linguistics, regiolect, bilingualism, language interference, Russian, Tatar

Проектируемый подкорпус русской речи билингвов, проживающих на территории Томской области является частью проекта «Лаборатории когнитивных исследований языка» филологического факультета Томского государственного университета по созданию корпуса «Томский региональный текст». Обоснованием разработки данного подкорпуса в рамках проекта является языковая специфичность Томского региона, которая тесно связана как с геополитическими, историческими процессами заселения этой территории представителями различных этносов, так и со смешанным, разнодиалектным составом русскоязычного населения [13. С. 33].

В предыдущих публикациях авторов проекта освещены отдельные аспекты концепции создаваемого корпуса – параметр региональности как базовый признак ограничения представленного материала [13], содержательное направление признаков сбалансированности и репрезентативности относительно данного типа корпуса [1], характеристика создаваемого корпуса по типологическим параметрам [16].

Работы по формированию лежащей в основе подкорпуса базы данных лексико-семантических и грамматических особенностей русской речи билингвов, проживающих на территории Томской области, выполняются коллективами исследователей Института русского языка им. В.В. Виноградова Российской академии наук и Томского государственного университета. Программная платформа корпуса разрабатывается под руководством проф. В.В. Поддубного.

Материалом для исследования послужили аудиозаписи бесед с двуязычными жителями Томска и Томской области, материалы бесед в открытых группах национальных объединений в социальных сетях, архивные документы. Поскольку подкорпус будет сбалансирован относительно устной речи, преимущество отдавалось сбору и обработке аудиозаписей.

Цель создания подкорпуса русской речи билингвов – отразить специфику русского языка в условиях его существования в билингвальных ситуациях. Своеобразие данных ситуаций заключается в языке контактирования, который обуславливает как

влияние на структуру русского языка, так и на речевую реализацию этого влияния. Поэтому нас интересует тип языка, который является контактирующим в каждой конкретной ситуации.

Проект подкорпуса русской речи билингов Томской области представляет пересечение двух актуальных направлений: изучения многоязычия и языковых контактов в рамках социолингвистики и создания лингвистических информационных ресурсов и баз данных в рамках корпусной лингвистики.

Явление би- и полилингвизма давно интересовало исследователей языка. В изучении языковых контактов раньше внимание обращали на взаимодействие двух систем языка, в фокусе внимания оказывались факты системного влияния: заимствование, конвергенция, дивергенция. В современной антропоцентричной лингвистической парадигме в фокусе внимания находится взаимовлияние языков в речи людей, владеющих двумя или несколькими языками.

Традицию и терминологию современных лингвистических изысканий в этой области заложил У. Вайнрайх в работах «Одноязычие и многоязычие» и «Языковые контакты», впервые изданных в 1953 и 1961 годах соответственно (переведены и изданы на русском в 1979 и 1972). Билингвизм (двуязычие) Вайнрайх определял как практику попеременного пользования двумя языками, а билингов – как лиц, осуществляющих такую практику [5, С. 22].

По охвату носителей, согласно классификации, предложенной Т.А. Бертагаевым в статье «Билингвизм и его разновидности в системе употребления», типы билингвизма, которые мы рассматриваем в данном проекте – индивидуальный и групповой (массовый) билингвизм [4, С.84-85].

Неотъемлемой частью речи билингва является интерференция. Вайнрайх определяет интерференцию как заимствование языковых элементов одного языка в другой и межъязыковое отождествление элементов двух языков и затем аналогичное по своей природе изменение элемента одного из этих языков под воздействием отождествленного с ним элемента другого языка [6, С. 31-32]. Л.И. Баранникова рассматривает интерференцию как частный случай взаимодействия языков, который может возникать лишь при регулярном использовании одним и тем же человеком или одним и тем же коллективом разных языков. При интерференции происходит изменение в структуре или элементах структуры одного языка под влиянием другого языка, причем не имеет значения, идет ли речь о родном, исконном для говорящего языке, или о втором языке усвоен-

ном позднее. [3, С. 88]. При этом заимствования и интерференция разграничиваются Баранниковой по объекту изменений: при заимствовании изменениям (ассимиляции) подвергаются те элементы одной системы, которые проникают в другую. При интерференции же изменениям подвергается сама заимствующая система, в ней появляются новые единицы, развиваются новые типы отношений между структурными элементами [3, С. 89-90]. Как подчеркивает Баранникова, большую роль в процессе интерференции играют именно те внутрисистемные отношения, которые никак не могут заимствоваться извне, поскольку в системе, испытывающей то или иное воздействие со стороны, развитие новых отношений возможно лишь при наличии внутренних тенденций, заложенных в самой данной системе. И выявление этих тенденций развития оказывается невозможным без влияния другой системы [3, С. 92]. В нашем исследовании мы рассматривали взаимовлияние русского и национальных языков с точки зрения Вайнрайха, поскольку невозможно рассматривать изменения семантических отношений, не принимая во внимание введения в лексику новых элементов, порождающих новые связи.

Явления интерференции охватывают все уровни языковой системы – фонетический, морфологический, семантический, синтаксический. Мы ограничиваем проблематику исследования рассмотрением обнаруженных лексико-семантических и грамматических явлений интерференции.

Ранее интерференцию в русском языке татар-билингвов, проживающих на территории Томской области, рассматривала в диссертации на звание кандидата филол. наук О.И. Гордеева. В работе Гордеевой освещаются фонетические и грамматические, но не лексико-семантические особенности влияния родного языка на усваиваемый [9, С. 17]. Часть выявленных ей закономерностей сохранилась в современном варианте русского языка томских татар, однако значительная часть явлений встречается реже, что обусловлено изменением языковой ситуации – после принятия в 1960х гг. законов о закрытии национальных школ и прекращении преподавания на национальном языке сфера употребления татарского языка резко сократилась. В результате на сегодняшний день татарским языком владеют и активно пользуются в основном представители старшего, реже – среднего поколения и те представители младшего поколения, которые заинтересованы в сохранении языка, однако, в последнем случае татарский язык чаще всего усваивается уже позже русского.

В области корпусной лингвистики тема региональных вариантов языка ещё не получила большого распространения. Один из наиболее близких идее Томского регионального корпуса проектов – корпус русскоязычных СМИ, функционирующих на территории Республики Беларусь, разработанный коллективами учёных из ИРЯ РАН (А.Я. Шайкевич, Г.И. Кустова, С.О. Савчук, Д.В. Сичинава) и Гродненского государственного университета им. Я. Купалы (Л.В. Рычкова, А. Станкевич и др.) [12]

### Грамматическая интерференция

«Грамматическая интерференция возникает тогда, когда правила расстановки, согласования, выбора или обязательного изменения грамматических единиц, входящие в систему одного языка, применяются к примерно таким же цепочкам элементов другого языка, что ведет к нарушению норм первого языка, либо тогда, когда правила, обязательные с точки зрения грамматики второго языка, не срабатывают в виду их отсутствия в грамматике первого языка» [5, С. 36].

Интерференция в области грамматики обусловлена различием грамматических категорий в разноструктурных языках. Следовательно, возможно предсказать области возникновения ошибок, вызванных интерференцией, охарактеризовав сходства и различия взаимодействующих языковых систем.

Мы рассматриваем следующие виды грамматической интерференции: а) грамматическая интерференция, вызванная расхождением в грамматических категориях (имен, глаголов и т.д.), б) грамматическая интерференция, вызванная расхождением в конструкции словосочетаний (управление глаголов, порядок слов в предложении и т.д.).

Поскольку в татарском языке категория рода отсутствует, русская категория рода для информантов выполняет только функцию формирования типов склонения. Этим объясняется частотность интерферентных явлений при выборе родового окончания или рода местоимений: кое-какое подвизка, У них однокомнатная была, они продали его. Неизменяемые имена собственные женского рода носители татарского языка вписывают в парадигму склонения существительных на –а/я: Марьяма зять (= зять Марьям), Марьяма дочь (= дочь Марьям). Некоторые ошибки в выборе категории рода характерны и для местных говоров: например, употребление мужского рода местоимения при объединении двух предметов женского рода: Две машины были, оба были УАЗик.

В области категории падежа наблюдаются как сходства, так и различия русского и татарского языка. Так, например, в татар-

ском существует отдельный притяжательный падеж (о категории принадлежности, которую он выражает, подробно будет сказано ниже), остальные падежи (основной, направительный, винительный, исходный, местно-временной) по функциям отчасти совпадают, отчасти отличаются от падежей русского языка.

На уровне согласования именных групп происходит употребление именительного падежа вместо косвенных в именных словосочетаниях: не учатся больше пять лет, Когда ребёнок появляется, года четыре-пять проходит...

Особого внимания заслуживают словосочетания, выражающие принадлежность. Поскольку наличие категории принадлежности – это ключевое отличие грамматической системы татарского языка, которое оказывает влияние так же и на лексику, факты интерференции, обусловленные перенесением этой категории и свойственных ей форм в русский язык, встречаются чаще других нарушений нормы, вызванных интерферентными явлениями в области грамматики словосочетаний.

Так, в архивных документах встречается передифференциация в сочетаниях, описывающих принадлежность: мнением Юсуповым вместо мнением Юсупа [15]. Для татарского языка характерны изафетные словосочетания, которые отличаются не присутствующим в русском языке типом синтаксической связи. В данном случае, когда имя существительное с аффиксом принадлежности третьего лица может определяться именем существительным также в притяжательном падеже, выражая тем самым принадлежность одного предмета другому, образуется изафетное словосочетание III типа [11, Т. 2. С. 35]. Для устной речи данное явление нехарактерно, что можно объяснить принадлежностью данных словосочетаний письменному стилю татарского языка [там же].

В области глагольных форм интерференция встречается не так часто, как в области именных, поскольку категории русского и татарского глагола во многом совпадают. Что касается нарушения норм литературного русского языка, то в русской речи томских татар встречаются формы глаголов, отличающиеся отсутствием частицы –ся: Вот он у меня проучил там сентябрь, октябрь, ... Однако, как отмечает Гордеева, колебания в употреблении частицы –ся с глаголами свойственны и окружающим русским говорам, и вообще говорам Томской области [9, С. 152], так что это может быть следствием влияния, скорее, окружающих диалектов, чем родного языка информантов.

Различия в управлении глаголов приводят к употреблению зависимых от них существительных с иными предлогами и в иных падежах, нежели предписывает норма русского литературного языка. Так, иногда русские наречия распознаются как существительные или прилагательные с предлогом, тогда приставка, распознанная как предлог, опускается, и «имя» используется в именительном падеже: Но там русские говорят татарски.

Синтаксическое влияние родного языка проявляется в русской речи томских татар чаще прочих интерференционных явлений. Поскольку татарский язык – постпозитивный, то в словосочетаниях зависимое слово чаще всего ставится перед главным, что и переносится на конструкции словосочетаний в русском языке: мне дали сразу направление, дали старшего техника звание, Марьяма зять. Соблюдение структуры SOP (структурная схема может быть расширена до Ats-S – Atv-O-P, где S – субъект, P – предикат, Ats – приименной атрибут (квалификатор), Atv – приглагольный атрибут, O – объект) [11, Т.3. С. 25] приводит к появлению подобных конструкций: Тот короче путь, Жизнь – сложная вообще штука, ...нас раздробить хотели...

Для русской речи томских татар в виртуальном пространстве характерны структуры типа “топик-комментарий”: а вообще активность:))) она пока отсутствует...:[18] Однако и в речи русской молодёжи можно встретить подобные структуры: А Европейский блок / он нам сейчас сильно нужен (Беседа в Новосибирске // Фонд «Общественное мнение», 2004) [14].

### **Лексико-семантическая интерференция**

Перейдём к рассмотрению фактов лексико-семантической интерференции. Поскольку под лексической интерференцией понимаются в большинстве случаев «все вызванные межъязыковыми связями изменения в составе лексического инвентаря, а также в функциях и употреблении лексико-семантических единиц, в их смысловой структуре» [10, С. 129], то можно выделить несколько уровней её рассмотрения. Так, С.В. Семчинский [17, С. 34] рассматривает 3 направления лексической интерференции:

- непосредственное заимствование лексических единиц;
- изменение структуры иноязычных единиц;
- заимствование значений и связей единиц с планом выражения из одного языка в другой.

Каждое из этих направлений интерференции присутствует в лексике томских татар, говорящих на русском языке. Рассмотрим их более подробно.

### **Непосредственные заимствования**

В результате анализа материала было обнаружено, что при высоком уровне владения обоими языками информанты в речи легко переключаются с одного на другой, и непосредственные заимствования редки. Однако определённая группа слов заимствуется в русскую речь из татарского языка. Это слова, обозначающие родственные или близкие отношения, употребляющиеся в бытовой речи в кругу семьи и друзей: – кызым ‘внучка’, бабай ‘дедушка’, эби ‘бабушка’, дуслар ‘друзья, товарищи’ и т.д. Они служат дискурсивными маркерами в семейном общении и употребляются даже молодым поколением татар, которое не владеет татарским языком.

Также определённые культурные слова могут быть включены в речь на русском языке, поскольку для них нет аналогий или имеющиеся аналоги являются неточными. Обычно это уже устоявшиеся заимствования: так, в русском устойчивом выражении не дай Бог производится замена лексемы Бог на Аллах. Кроме того, часто заимствуются этикетные формулы: салам ‘здравствуйте’, рахим итэгез ‘добро пожаловать’.

Некоторые различия в грамматических категориях оказывают влияние и на лексику. Так, поскольку в татарском языке категория принадлежности выражается синтетически, при помощи падежных окончаний, нет потребности в выражении принадлежности при помощи притяжательных местоимений. Это приводит к ошибкам в употреблении притяжательных местоимений в русской речи томских татар: У них однокомнатная была, они продали его. Я спросил у него, сколько тебе осталось плавать, говорит, еще 5 лет и будет своей.

### **Изменение структуры иноязычных единиц**

Факты интерферентного влияния на структуру слов в русском языке томских татар единичны, в основном они присутствуют в речи информантов, несвободно владеющих русским языком и испытывающих сложности с кодовым переключением. Наиболее ярким примером является слово послевчера (= позавчера).

Иногда встречаются модификации слов национального языка, использующихся в русской речи – например, добавление русского диминутивного суффикса в сообщениях из социальных сетей: дусларчики (тат. дуслар – ‘друзья, товарищи’).

### **Заимствование значений и связей единиц с планом выражения из одного языка в другой**

Наиболее сложным представляется явление заимствований значений и связей

единиц с планом выражения из родного языка говорящего в неродной. Определить наличие интерференции подобного рода можно в случаях неточного или неудачного словоупотребления или же нарушения лексической сочетаемости.

Лексическая сочетаемость Ю.Д. Апресяном определяется как допустимость одних лексем в качестве распространителей определённого слова и недопустимость других, без нарушения семантических запретов: так, возможно сочетание сбрасывать (в значении 'резко уменьшать') скорость, но не сбрасывать расходы [2, С. 61].

В русской речи билингов нарушение лексической сочетаемости встречается при неверном выборе многозначных слов для передачи устойчивых выражений. Восстановить лексически и стилистически верный вариант при подобной интерференции бывает непросто, поскольку во фразе может быть допущено несколько ошибок подобного рода. Например, в предложении Однако, в большинстве своем, танцам в Татарстане характерны светлые и добрые замыслы. вместо лексемы характерны следовало употребить лексему свойственны, а вместо лексемы замыслы – лексему намерения, однако даже в этом случае сохранилась бы лексическая ошибка, поскольку существительные, не обозначающие человека, не могут сочетаться с абстрактными существительными, имеющими сему 'продукт ментальной деятельности человека'. Причиной употребления именно лексемы замыслы вместо намерения послужила многозначность татарского слова ният, обозначающего и 'намерение, цель', и 'замысел, умысел' [7. С. 167; 8. С. 222].

Часто сходная семантика и внутренняя форма слов русского языка также становится причиной замены слов в устойчивых выражениях. Так, в репликах Где он? Какой он переводчик? Ну так/ для обзора. и Сразу взять и работать тебя не будет/ надо как-то обзорно смотреть...лексемы обзор и обзорно употреблены вместо лексем пример и примерно.

Таким образом, на материале русско-татарского билингвизма были выявлены лексико-семантические и грамматические особенности речи билингов, отличающие порождённые ими тексты от текстов, порождаемых монолингвами. Для дополнения базы данных планируется описание различных остальных языковых систем, которые взаимодействуют на территории Томской области и определение необходимого количества текстов для каждой билингвальной ситуации с целью внесения их в подкорпус русской речи билингов с соблюдением принципов

сбалансированности и репрезентативности. Также планируется разработка принципов лингвистической разметки подкорпуса русской речи билингов, проживающих на территории Томска и Томской области.

При дальнейшем развитии проекта на его базе возможны проведения исследовательской языковой ситуации в Томской области, исследований, посвящённых национальной идентичности, этнической специфике репрезентации языковой личности.

*Публикация подготовлена в рамках под-держанного РФФИ научного проекта № 14-36-50404 «Создание базы данных лексико-семантических и грамматических особенностей русской речи билингов, проживающих на территории Томской области».*

#### Список литературы

1. Sologub O., Rezanova Z., Temnikova I. The Concept of the Tomsk Regional Corpus: Balance and Representativeness // The XXV annual international academic conference, Language and culture, 20-22 October 2014 / Procedia – Social and Behavioral Sciences, 154 (2014) pp.175-178.
2. Апресян Ю.Д. Лексическая семантика. – М., 1974.
3. Баранникова, Л.И. Сущность интерференции и специфика её проявления / Проблемы двуязычия и многоязычия – М.: Наука, 1972. – с. 88-98.
4. Бертагаев Т.А. Билингвизм и его разновидности в системе употребления // Проблемы двуязычия и многоязычия. – М.: Наука, 1972 – с. 82-88.
5. Вайнрайх У. Одноязычие и многоязычие // Новое в лингвистике. Вып. VI. Языковые контакты. – М., 1972. – С. 25-60.
6. Вайнрайх У. Языковые контакты. – Киев, 1979.
7. Ганиев Ф.А. Русско-татарский словарь. – М.: ИН-САН, 1997.
8. Ганиев Ф.А. Татарско-русский словарь. –Казань: Татарское книжное издательство, 1988.
9. Гордеева О.И. Некоторые закономерности влияния родного языка на усваиваемый язык в процессе становления двуязычия (на материале русского и татарского сибирских говоров): дисс. канд. филол. наук. – Томск, 1965.
10. Жлуктенко Ю.А. Лингвистические аспекты двуязычия. – Киев: Вища школа, 1974.
11. Закиев М. З. Татарская грамматика в 3 т. – Казань, 1995.
12. Кустова Г.И., Савчук С.О. Изучение лексико-семантической и социокультурной специфики русской речи на территории Республики Беларусь (на материале текстов СМИ) // Труды междунар. конф. «Корпусная лингвистика-2013». – СПб.: С.-Петербургский гос. университет, Филологический факультет. – 2013. – С. 344-352.
13. Мишанкина Н.А. Лингвистический корпус «Томский региональный текст»: теоретико-методологическое обоснование проекта // Вестн. Том. гос. ун-та. – 2014. – № 389. – С. 28–37.
14. Национальный корпус русского языка. – URL: <http://guscorpora.ru>.
15. ОГКУ «Центр новейшей документации истории Томской области», Ф. № 1, О. № 1, Д. № 1533, Л. № 98.
16. Резанова З.И. «Томский региональный корпус»: типологически релевантные параметры сбалансированности и репрезентативности // Вестн. Том. гос. ун-та. Филология. – 2015. – №1 (33). – С. 38-50.
17. Семчинский, С.В. Семантическая интерференция языков (на материале славяно-восточнороманских языковых контактов): автореф. дис. ... д-ра филол. наук / С.В. Семчинский. – Киев, 1973.
18. Томские татары. – группа в социальной сети «Вконтакте». URL: <http://vk.com/club27426566>.