

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра теории, истории языка и прикладной лингвистики

**НОМИНАЦИИ ЛИЦ В ДИАЛЕКТНЫХ ТЕКСТАХ
И ПРОБЛЕМЫ ДИАЛЕКТНЫХ КОРПУСОВ**


АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студентки 4 курса 441 группы
направления 45.03.03 Фундаментальная и прикладная лингвистика
Института филологии и журналистики
Силенко Алёны Андреевны

Научный руководитель

д.ф.н., профессор

14.06.16

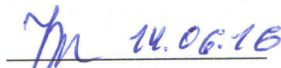

подпись, дата

В.Е. Гольдин

инициалы, фамилия

Зав. кафедрой

д.ф.н., профессор


подпись, дата

О.Ю. Крючкова

инициалы, фамилия

Саратов, 2016

ВВЕДЕНИЕ

Данная работа состоит из двух частей; первая («Параметры организации диалектных корпусов») представляет собой сравнение доступных для рассмотрения диалектных корпусов, в том числе корпусов, находящихся в стадии разработки; вторая («Система номинаций лиц в говоре (мегорский говор)») – посвящена специфике системы номинации лиц в диалектных текстах как явлению, которое может получить корпусное воплощение.

Общая цель данной работы - рассмотреть современное положение диалектной корпусной лингвистики как части компьютерной лингвистики и технические возможности существующих и создающихся текстовых диалектных корпусов, понять, как и на каком фоне создается СарДК, выявить состав текстовых номинаций лиц в диалектной речи и определить их текстовое функционирование.

Объект исследования – записи диалектной речи, сделанные в с. Мегра Вытегорского района Вологодской области, а также статьи, посвящённые теоретическим разработкам диалектных корпусов.

Предмет исследования – специфика системы номинаций лиц в диалектной речи (корпусный аспект).

Актуальность и новизна данной работы, как нам представляется, обусловлена исследованием диалектного материала с позиций корпусной лингвистики. С одной стороны, это палитра именованных лиц и предметов в диалекте. С другой стороны, – реализация этой системы в корпусах и оценка технических возможностей создающихся текстовых диалектных корпусов.

Создание текстовых корпусов – одно из приоритетных направлений современной лингвистики. В настоящее время это направление выделилось в отдельную отрасль лингвистики – корпусную лингвистику. Существуют корпусы письменной, устной, устной и письменной речи. Но на данный момент существует острая нехватка диалектных корпусов. Хотя сегодня они создаются во многих странах, большая часть их не завершена или завершена, но ещё не от-

крыта пользователям. Доступнее оказываются теоретические корпусные разработки и предварительные описания полученных авторами результатов в пилотных исследованиях. Поэтому в большинстве случаев нам приходилось работать с теоретическими статьями, в которых описываются характер строения и поисковые возможности создающихся диалектных корпусов. Проблемам корпусной диалектологии посвящена **первая глава** работы.

Во **второй главе** рассматриваются номинации лиц в диалектных текстах, подготовленных к представлению в Саратовском мультимедийном диалектном корпусе (СарДК). Из собрания использованы расшифровки аудиозаписей пяти информантов: Л.Ф. Битяковой (1929 г.р.), М.К. Власова (1923 г.р.), С.М. и Г.С. Копачёвых (1928 г.р. и 1931 г.р.), А.Г. Митина (1914 г.р.). Общий объём расшифрованных аудиозаписей – 61800 словоупотреблений, из которых около 33 тыс. словоупотреблений принадлежат лицам мужского пола и около 28 тыс. словоупотреблений – лицам женского пола. Система диалектных номинаций лиц рассматривается в главе как отражение картины мира диалектоносителя.

Отдельный параграф главы посвящен рассмотрению номинаций лиц с точки зрения позиций, занимаемых этими номинациями в строении текста. Теоретической основой данного анализа является предложенная Н.Д. Арутюновой типология текстовых функций номинаций лиц.

В **Заключении** подводятся итоги работы.

Библиографический список состоит из 38 источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе данной работы обсуждаются вопросы, связанные с параметрами организации диалектных корпусов.

Поскольку создание и использование диалектных корпусов – это достаточно молодое научное направление, на данный момент ещё не сложилась единая концепция диалектных текстовых корпусов. Поэтому на данном этапе важно обсуждение основных параметров строения диалектных корпусов.

Наша попытка дать обзор ведущихся сегодня работ по созданию диалектных текстовых корпусов строится на доступных в настоящее время материалах по Корпусу грузинских диалектов, Скандинавскому диалектному корпусу, Цифровому корпусу болгарских диалектов, Эстонскому диалектному корпусу, Корпусу вепсского языка, Польскому диалектному корпусу, корпусу Ustja River Basin Corpus, разрабатываемому в Высшей школе экономики (Москва) и Саратовскому мультимедийному текстовому диалектному корпусу (СарДК). Полнее других нами рассмотрены материалы Хельсинкского корпуса британских диалектов, Корпуса вепсского языка, Польского диалектного корпуса и Ustja River Basin Corpus, поскольку это одни из немногих современных диалектных корпусов, представленных в свободном доступе в сети Интернет.

Заметим, что в данной работе не рассматривается диалектный подкорпус НКРЯ по тем соображениям, что «говоры как реально функционирующие системы в НКРЯ не рассматриваются» [Летучий 2005: 229]. Нас же интересует полноценный репрезентативный диалектный корпус, который наглядно отражает все особенности диалектной речи при коммуникации.

Сравнение проводилось по перечню параметров, предложенному О.Ю. Крючковой и В.Е. Гольдиным¹:

1) принципы отбора диалектного материала и критерии репрезентативности диалектного корпуса;

¹ Гольдин, В.Е., Крючкова О.Ю. Корпус русской диалектной речи: концепция и параметры оценки // Компьютерная лингвистика и интеллектуальные технологии : по материалам ежегодной Международной конференции «Диалог». Вып. 10 (17). М.: РГГУ, 2011. С. 359-367

- 2) принципы членения речевого континуума в корпусе;
- 3) параметры выдачи текстовых фрагментов;
- 4) формы представления диалектных текстов в корпусе;
- 5) правила и виды аннотирования текстовой базы корпуса;
- 6) параметры метаразметки диалектных корпусов;
- 7) представление в диалектном корпусе нелингвистической информации;
- 8) оптимальные для диалектологических исследований возможности.

Приведём несколько примеров диалектных корпусов по каждому параметру.

Каждый из рассмотренных нами корпусов претендует на статус источника лингвистических и лингвокультурологических исследований, то их можно назвать научными источниками нового типа, соответствующими общей идеологии корпусной лингвистики, поскольку диалектные тексты могут стать доступными пользователям и исследователям преимущественно только в случае включения их в корпуса.

Принципы членения речевого континуума в корпусе. Данный параметр напрямую связан с целями каждого конкретного диалектологического корпуса. Так, создатели Корпуса грузинских диалектов решили объединить в корпусе диалектные словари и текстовые материалы, т.к., по мнению разработчиков, это гарантирует полное представление слов в корпусе и дает дополнительный контекст для понимания лексем (См.: [Beridze, Nadaraia 2009]).

В СарДК единицей членения речевого потока является полноценный текст, который соответствует зафиксированному аудиоаппаратурой участку непрерывного общения (См.: [Гольдин, Крючкова: 2011]).

О параметре выдачи текстовых фрагментов в Хельсинкском корпусе британских диалектов сказано, что в настоящее время – это минимальный контекст (3-4 строки, и, как правило, это первые предложения в записи). Полный вариант записи можно скачать (См.: [URL: <http://www.helsinki.fi/varieng/CoRD/corpora/Dialects/>]).

В Вепском корпусе аудиозаписи и письменные тексты разведены и никак не пересекаются (См.: [URL: <http://vepsian.krc.karelia.ru/about/>]). Все остальные корпуса предлагают поиск в тексте интересующих слов, но не сами тексты.

Если говорить о **формах представления текстов в корпусе**, то здесь мы вновь с большей уверенностью можем сказать о Хельсинкском корпусе британских диалектов, Вепском диалектном корпусе и портале Ustja River Basin Corpus. Диалектные тексты представлены только в виде орфографической записи (См.: [URL: <http://www.slavist.de/Pushkino/index.php>]).

В Хельсинкском корпусе британских диалектов и Вепском диалектном корпусе все высказывания информантов записаны как можно тщательнее, в том числе со всеми колебаниями, повторами, т.е. текст «старается» отразить всё то, что обычно не представлено в письменной речи, но характерно для устной (См.: [URL: <http://www.helsinki.fi/varieng/CoRD/corpora/Dialects/>]).

Виды и правила аннотирования текстовой базы корпуса. В Эстонском диалектном корпусе записи сопровождаются фонетически расшифрованным текстом, файлом с упрощенной транскрипцией и описанием.

Стоит заметить, что при всей новизне портала Ustja River Basin Corpus, в нём не решены вопросы метаразметки текстов: она отсутствует. Однако можно просмотреть частотные токены, словоформы, количество токенов каждого информанта и список словоформ, у которых не зафиксировано начальной формы. Все эти списки расположены в порядке убывания (См.: [URL: <http://www.slavist.de/Pushkino/index.php>]).

Параметры метаразметки диалектных текстов. По данному параметру корпуса оказываются схожими. Во всех корпусах можно найти информацию о диалектоносителе. В одних корпусах эта информация оказывается более подробной, в других - менее подробной. Так, в Корпусе грузинских диалектов выделяется информация о родителях диалектоносителя, их национальность, информация о миграциях диалектоносителя и его семьи, год миграции, кто именно мигрировал и т.д. (См.: [Beridze, Nadaraia 2009]).

В СарДК помимо основных данных (указание места, времени записи, темы, объёма текста) включены также сведения о конкретной ситуации общения, об адресатах речи, о времени описываемых в общении событий и др. (См.: [Гольдин, Крючкова: 2011]).

Оптимальные для диалектологических исследований возможности. Все рассмотренные корпуса (за исключением Хельсинкского и Вепского) в настоящее время находятся на стадии разработки, поэтому постоянно уточняются общие принципы их построения. И поэтому широкое обсуждение данных проблем является актуальным для корпусной диалектологии.

Таким образом, можно утверждать, что можно утверждать, что, несмотря на некоторые достижения в создании диалектных корпусов, глубоко аннотированный репрезентативный текстовый диалектный материал ещё мало доступен широкому кругу пользователей, поэтому теоретическая разработка проблемы и создание диалектных корпусов является одной из самых важных задач корпусной лингвистики.

В первой части второй главы проводится сравнительный анализ систем номинаций лиц литературного языка и мегорского говора с. Мегра Вологодской области, который относится к Онежской межзональной группе говоров северного наречия. Материалами исследования являются Русский семантический словарь под ред. Н.Ю. Шведовой и расшифрованные тексты аудиозаписей диалектоносителей данного говора.

Задача заключается в том, чтобы установить специфику номинаций данного диалекта на основе полученных группировок.

Для группировки и общей оценки номинаций лиц в говоре мы соотнесли диалектные номинации лиц с классификацией, представленной в Русском семантическом словаре под ред. Н.Ю. Шведовой.

В ходе исследования были выявлены 4 основные группы номинаций лиц, которые представлены в речи всех диалектоносителей:

- Собственно обращения;
- Номинации родственников;

- Номинации профессий, видов деятельности;
- Собирательные номинации и совокупности лиц.

Каждая из данных групп, в свою очередь делится на подгруппы: группа «Собственно обращения» включает в себя подгруппу «Обращения с эмоциональным компонентом» (*бедненькие, дурак*), «Фамильярные, дружеские обращения» (*любошка, дорогуша*), группа «Номинации профессий, видов деятельности» включает подгруппы «Номинации должностей» (*председатель колхоза, секретарь райкома*), «Номинации лиц по увлечениям» (*гармонисты*), «Номинации родовых понятий, связанных с той или иной деятельностью» (*мастер, специалист*) и т.д., группа «Собирательные номинации и совокупности лиц» - «Номинации по полу» (*мужики женщины*), «Номинации по возрасту» (*молодёжь, старые, девчата*), «Характеризующие номинации» (*труженики, безработные*).

Всего в рассмотренных текстах нами было выявлено в совокупности 735 употреблений номинаций лиц.

Для диалектоносителей женского пола оказываются наиболее важными упоминания своих близких родственников и обращения к кому-либо (к диалектологу или какому-либо лицу, о котором идёт повествование). Это проявляется в относительной частотности данных номинаций. В то время как для диалектоносителей мужского пола важными и потому частотными оказываются номинации лиц, связанные с профессиями и различными социальными группами людей.

Имена нарицательные более прямо отражают диалектную картину мира, показывая, что является наиболее важным для носителей данного говора. Выявление номинаций лиц различными группами слов (названия лиц по профессии, по полу и т.д.) раскрывает специфику взаимоотношений диалектоносителей. В определённых ситуациях общения для диалектоносителей важнее выделить человека через его профессию, а неродственные связи или через указание его должности, а не возраста и т.д.

Специфика системы номинаций лиц данного диалекта раскрывается в сопоставлении полученных групп у мужчин и женщин диалектоносителей и сравнении этих групп с системой номинаций, представленной в Русском семантическом словаре. Существуют различия в употреблении определённых номинаций у мужчин и женщин. Были выявлены подгруппы, которые характерны только для диалектоносителей мужского пола или количественно преобладают у них, а также подгруппы, которые количественно равны у мужчин и женщин. Это свидетельствует о традиционном разделении обязанностей у женщин и мужчин в сельской местности.

Целью **второго раздела второй главы** является исследование текстовых функций номинаций лиц в диалектной речи на материале мегорского говора. В качестве инструмента анализа нами использована референтная модель номинаций лиц, предложенная Н.Д. Арутюновой [Арутюнова 1977]. Заметим, что применение данной методики по отношению к исследованию диалектной системы номинации лиц ранее не зафиксировано.

Номинация лица зависит от роли ее в тексте. Н.Д. Арутюнова выделяет следующие текстовые позиции, которые оказывают влияние на номинации лиц (См.: [Арутюнова 1977]):

- экзистенциальная (интродуктивная) – позиция введения персонажа в повествование – 212 словоупотреблений:
 - ❖ *моя сестра родная/ тётя Саши Некипелова/ самая старшая у нас// не дождала девяноста двух лет//* (Власов М.К.);
 - ❖ *у меня родители неграмотные были* (Копачёва Г.С.);
- идентифицирующая – позиция отождествления упоминаемого лица с лицом, уже введенным в повествование – 77 словоупотреблений:
 - ❖ *жена местная ходила проверяла налоги/ она была налоговым инспектором/ достался ей мой участок//* (Власов М.К.);
- позиция предиката – 55 словоупотреблений:

❖ *всё отобрали/ сделали нищима// у нищих вот этих у колхозников да у... заво... фабрики заводы закрывают// а шт... кто богатеет-то?/ кто там у власти/ наверно/ так?// они стали миллионеры тут//* (Битякова Л.Ф.);

- позиция обращения – 87 словоупотреблений:

❖ *ой/ дорогуша/ отца я не помню/ я родился при нём ещё//* (Митин А.Г.);

❖ *погодите любюшка/ я/ не/ не понимаю ничего что ты говоришь//* (Копачёва Г.С.)

Информанты зачастую не осознают тот факт, что диалектолог не понимает, о ком идёт речь при коммуникации. Также для диалектоносителей характерно упоминание тех лиц, которые могут быть как-то связаны с ситуацией общения, но не являются при этом её главным участником или центром коммуникации. В таких случаях встречаются примеры номинаций лиц, которые мы называем **первой номинацией в тексте** (316 словоупотреблений), поскольку такую номинацию нельзя отнести к интродуктивной, но и нельзя оставить без внимания, например:

- *нао денег было заработать/ и Алексей как раз учился на механика/ в Черезповце//* (Битякова Л.Ф.);

- *много сейчас церковей строят/ но у нас таких нету энтузиастов/ никто наверно не хочет никто// там поклонный крест вон Серафимовна это... поставила/ Валентина Серафимовна/ попросила крест/ сделали/ вот//* (Битякова Л.Ф.);

Экзистенциальная номинация выбирается в соответствии с той областью бытия, в которую она вводится. Она указывает на общие признаки, свойств лица, о котором идёт речь, т.е. причисляет его к достаточно широкому классу лиц. Диалектоносители много рассказывают о своей деревне, семье, употребляя интродуктивную номинацию. Также мы выделили такую особенность, как **первая номинация в тексте**, которая по количественным данным существенно превосходит другие типы номинаций. Это объясняется восприятием диалектоносителем описываемой ситуации единым целым, которое не

будет являться таковым без упоминания связанных с этим целым, пусть даже незначительных и вскользь упоминаемых, лиц.

Идентифицирующая номинация выражает индивидуальные признаки лица, которые выделяют это лицо в данном классе. Как правило, выбор идентифицирующей номинации регулируется прагматическими факторами. Небольшое количество идентифицирующих номинаций – это особенность общения на диалекте.

Предикатная номинация направлена на раскрытие существенных черт лица. Как нам кажется, диалектоносители при помощи таких номинаций имеют возможность более точно выразить своё мнение относительно упоминаемого лица. По встречаемости в диалектной речи предикатная номинация находится на последнем месте. Это лишний раз подтверждает нашу гипотезу о том, что упоминаемые лица рассматриваются в тесной связи с ситуацией-темой, которая, в свою очередь, выходит для диалектоносителей на первый план.

При **обращении** часто используются имена релятивной семантики, фиксирующие отношения между собеседниками. Выбор такого имени определяется, как правило, взаимоотношениями собеседников. Поэтому номинаций-обращений в речи диалектоносителей достаточно много.

По нашему мнению, необходимо выделить специальный раздел для диалектных номинаций лиц в метатекстовой разметке диалектного корпуса. Необходимо указывать адрес, в котором встречается та или иная номинация, количество их употреблений в конкретном тексте, использование данной номинации в качестве обращений, наличие эмоционального компонента. А также нам кажется важным и целесообразным рассмотреть номинации одних и тех же лиц, которые встречаются в речи разных диалектоносителей одного села или деревни и которые по-разному ими именуется. Это позволит исследовать, какие именованные используются для лиц в речи, что на это влияет и т.д.