

ЛИНГВИСТИКА

Ю. Д. Абаева

ДИАЛЕКТНЫЙ КОРПУС КАК СРЕДСТВО СОХРАНЕНИЯ И ИЗУЧЕНИЯ БУРЯТСКОЙ ПРОСОДИИ¹

Исследование посвящено вопросам разработки Бурятского диалектного корпуса, ориентированного на исследование просодии бурятских диалектов. В работе отмечается многодиалектный характер бурятского языка, анализируется современное состояние диалектов, описываются основные труды по бурятской диалектологии. В результате автор приходит к выводу, что, хотя диалекты активно функционируют, их нынешнее состояние вызывает опасение. Вместе с тем в 80-х гг. прошлого века была свернута программа по изучению диалектов и по сей день отсутствуют масштабные, планомерные исследования. Создание такого ресурса обусловлено все усложняющейся социолингвистической ситуацией с бурятскими говорами, а также отсутствием в бурятской диалектологии современных исследований вообще и по просодии в частности.

В работе описываются основные этапы работы над корпусом, требования к сбору материала, принципы его сегментирования и аннотирования. Материал записывался в районах этнической Бурятии, дикторы — мужчины и женщины, с детства владеющие своим диалектом и использующие его в качестве средства повседневного общения. Программа записи состоит из двух блоков: подготовленные фразы для исследования интонации и спонтанная повседневная речь. Анализ интонации проводится при помощи программы PRAAT. Для сегментации и аннотирования используется программа ELAN.

В работе особое внимание уделяется принципам просодической разметки бурятской речи, используемой при аннотировании корпуса. В результате анализа имеющихся просодических транскрипций были разработаны основные принципы просодического аннотирования. В данном диалектном корпусе бурятского языка просодическая транскрипция включает сегментацию на речевые отрезки разной степени завершенности и коммуникативной направленности, указание на место интонационного центра и участки просодической выделенности, интонационный контур речевого отрезка. Разработаны знаки для условных обозначений.

Работа по созданию Диалектного корпуса находится в начале пути, намечаются планы и дальнейшие перспективы работы. Данный ресурс будет полезен не только для исследования диалектной интонации, но и для бурятской диалектологии в целом.

Ключевые слова: бурятские диалекты, звуковой диалектный корпус, просодия, ударение, интонация, просодическая аннотация.

Введение

На территории Российской Федерации носители бурятского языка компактно проживают в трех регионах: Республике Бурятия, Иркутской области (Усть-Ордынский округ) и в Забайкальском крае (Агинский округ). За рубежом буряты проживают в Китае и Монголии. Как видим, территория их расселения достаточно обширна. Если обратиться к истории формирования бурятского этноса, то увидим, что он формировался из достаточно большого количества крупных и малых племенных групп, в результате их постоянной миграции и взаимодействия с другими народами (племенами тунгусов и тюрков и др.). Все эти факторы обусловили значительную вариативность бурятского языка, его многодиалектный характер.

¹ Статья подготовлена в рамках государственного задания (проект XII.193.1.5 «Ментальность монгольских народов в зеркале языка», № АААА-А17-117021310266-8).

Традиционно бурятские диалекты РФ делятся по территориальному признаку (на западные, восточные и южные), либо по родо-племенной принадлежности, что является более точным критерием. Согласно последнему критерию можно выделить четыре больших диалектных группы (наречия): хоринская, эхирит-булагатская, аларо-тункинская, цонголо-сартульская, каждая из которых включает несколько родственных говоров. Помимо этих больших групп имеется еще два говора, стоящих особняком: нижнеудинский (самый западный) и говор ононских хамниган (Бураев, 1996: 14). Что касается бурятских диалектов Монголии и Китая, то их нужно выделить в отдельные группы, так как, к примеру, шэнэхэнские буряты, проживающие на севере КНР, хотя и являются в основном хоринцами, но их говор имеет свои специфические черты, отличающие его от языка хоринских бурят РФ. То же самое касается и бурятского языка в Монголии.

Относительно современного состояния бурятских диалектов можно сказать, что они живы и активно функционируют. В сельской местности имеются населенные пункты, где повседневное общение ведется на бурятском языке, в том числе среди молодежи и детей. Если литературный бурятский язык более характерен для СМИ, образования, культуры, то разговорная речь — это чаще всего речь диалектная.

О том, что диалекты активно функционируют, свидетельствует, к примеру, популярность среди бурят радио Buryad FM, специализирующегося на продвижении бурятской диалектной речи, пропагандирующего сохранение и поддержание диалектных различий. Некоторое время назад выпускалась газета «Нютаг хэлэн / Диалекты», которая также, выступала площадкой для носителей диалектов и пользовалась большим интересом у читателей (Баторов, 2016).

Однако лингвистическая ситуация по диалектам складывается неравномерно. Если говорить о бурятском языке на территории РФ, то в наиболее трудном положении находится нижнеудинский диалект, возможно в силу своей территориальной отдаленности и изолированности от других бурятских говоров. Лингвистической экспедицией 2016 года в Нижнеудинский район Иркутской области было выявлено, что численность носителей этого диалекта составляет немногим более ста человек в с. Кушун и четыре семьи в с. Мунгутубулук. Старшее поколение еще помнит родной язык, усвоенный в детстве, хотя и редко пользуется им в повседневном общении, а молодежь и дети его практически совсем не знают.

Более благоприятные условия для сохранения и развития диалектов складываются на территории бурятских округов и в Республике Бурятия, где диалекты находятся в близком соседстве друг с другом. Обычно в пример по уровню владения бурятским языком приводится Агинский Бурятский округ Забайкальского края, где жители говорят большей частью на родном языке, обучая ему своих детей.

Тем не менее современные процессы урбанизации и глобализации делают свое дело и даже, казалось бы, в благополучных в языковом плане регионах ситуация меняется не в лучшую сторону. Хотя бурятские диалекты на сегодняшний день достаточно активно используются, нельзя не отметить тот печальный факт, что снижается количество детей, говорящих на родном языке, а это значит, что язык, пока активно функционирующий, находится перед реальной опасностью быть забытым. В связи с этим, создание диалектного корпуса бурятского языка, как ресурса для сохранения и изучения бурятских говоров, представляется весьма актуальным.

Наиболее активный и продуктивный период исследования бурятских диалектов — это середина прошлого века, когда осуществлялись диалектологические экспедиции в различные регионы этнической Бурятии и велся активный сбор языкового материала. Группы диалектологов жили рядом с информантами и вели наблюдения за их повседневной речью, образом жизни, фиксировали не только языковые особенности, но и иную информацию (этнографические заметки, фольклор, исторические предания и др.). По результатам таких выездов были опубликованы работы, в которых достаточно подробно описывались диалектные характе-

ристики на всех языковых уровнях, например «Исследование бурятских говоров» в двух выпусках (Исследование ..., 1965, 1968). Из более поздних диалектологических работ следует отметить труды по нижнеудинскому говору (Рассадин, 1999), так называемой присаянской группе говоров (Рассадин, 1996), по исторической фонетике монгольских языков (Рассадин, 1982), исследование консонантизма хоринского диалекта (Соктоева, 1988), диахроническое исследование бурятских диалектов (Будаев, 1992) и др.

На сегодняшний день таких масштабных и планомерных исследований бурятских говоров не проводится. Имеются отдельные статьи, посвященные характеристикам того или иного говора, например: Н. Г. Замбулаева (2011), В. Д. Патаева (2009), Ц. Ц. Бальжинимаева (2012), В. И. Семенова (2013) и др. Как пишет Г. А. Дырхеева: «... в связи с общеизвестными переменами в стране фактически не уделялось внимание особенностям трансформации диалектов и говоров, их месту в языковой ситуации региона в период 80-х гг. XX в. — начала XXI в. Отсутствие финансирования вынудило свернуть программы по их изучению, и оказалось практически невозможно некоторые из них восстановить» (Дырхеева, 2017: 80).

В диалектологических работах прошлого века наибольшее внимание уделялось сегментной фонетике и лексическому составу как языковым уровням, в которых наиболее ярко проявляются диалектные отличия. Просодические характеристики не описывались, имеются лишь некоторые заметки относительно ударения или интонации, основанные на перцептивном восприятии.

Исследование суперсегментных характеристик бурятского языка началось в период развития экспериментально-фонетических методов. Однако в основном исследования проводились на материале литературного бурятского языка. Диалектная интонация до сих пор в должной мере не исследована.

В связи с вышеизложенным возникла идея создания звучащего корпуса бурятских диалектов, нацеленного главным образом на исследование суперсегментных характеристик современной бурятской речи. На сегодняшний день корпусный метод представляется наиболее оптимальным средством исследования диалектной интонации, который не только позволяет включить в анализ большой массив данных, но и является средством фиксации и сохранения языкового материала.

При разработке концепции нашего диалектного корпуса, его наполнения и принципов составления мы опирались на опыт уже имеющихся звучащих корпусов. Диалектные корпуса создаются для большинства языков мира, например, английского (Survey of English Dialects), скандинавских языков (The Nordic Dialect Corpus), болгарского (Bulgarian Dialectology as Living Tradition), русского (Диалектный корпус в составе НКРЯ, Кубанский диалектный корпус, Томский диалектный корпус и др.). Имеются разработки по монгольским языкам: в Национальный корпус калмыцкого языка включены записи устной речи калмыков старшего поколения, с целью выявления диалектных элементов (Куканова, 2013). Во Внутренней Монголии коллективом ученых под руководством Д. Батора разработан масштабный корпусный проект по монгольским языкам и диалектам Внутренней Монголии (Китай), Монголии и России (Russian news).

В мировой лингвистике имеется достаточный опыт создания корпусов, нацеленных на изучение просодических характеристик, в том числе на диалектном материале. Например: Русский интонационный корпус (Архипов, 2012), Интонационная вариативность в английском (The IViE Corpus), содержащий информацию об интонации диалектов английского языка, Звуковой корпус для изучения просодии младописьменных языков на примере диалектов ненецкого языка (Шерстинова, 2006) и др.

Таким образом, актуальность создания Бурятского диалектного корпуса обусловлена с одной стороны все усложняющимся положением бурятских диалектов, с другой — отсутствием современных всеобъемлющих исследований по бурятской диалектологии вообще и по просодии в частности. Планируемый диалектный корпус будет нацелен на фиксацию звуча-

щей стороны современных бурятских диалектов и в первую очередь их просодических характеристик. Звуковой корпус позволит сохранить произносительные особенности живой бурятской разговорной речи и станет основой для их изучения. Однако исследовательские возможности корпуса не будут ограничиваться только суперсегментными характеристиками, транскрибированные и аннотированные тексты позволят вести исследование диалектов на всех языковых уровнях. Бурятский диалектный материал может войти в качестве диалектного подкорпуса в активно создаваемый ныне Корпус бурятского языка (Бурятский корпус).

Этапы работы над корпусом

На сегодняшний день ведутся подготовительные работы по сбору и систематизации записей бурятской диалектной речи, разрабатываются принципы их оформления в виде звучащего корпуса. Основные этапы работы над корпусом выглядят следующим образом:

- 1) сбор материала, в результате которого получаем массив звуковых файлов для последующей обработки;
- 2) паспортизация полученных записей: таблицы с информацией об информанте, времени и месте записи, длительность звучания и т. д. Присвоение номеров полученным звуковым файлам;
- 3) аудирование носителями языка;
- 4) аннотирование в программе ELAN: расшифровка в орфографической записи с сегментацией на синтагмы, слова, просодическая аннотация;
- 6) измерение акустических параметров в программе PRAAT (тон, интенсивность, длительность);
- 7) лингвистическая интерпретация полученных данных, описание диалектной интонации.

Материал исследования

Речевой материал для корпуса записывался в местах проживания носителей диалектов на территории РФ: в Республике Бурятия, Иркутской области и Забайкальском крае. Имеются записи по бурятским диалектам Монголии (северные районы) и КНР (Автономный район Внутренняя Монголия). Информантами выступали как мужчины, так и женщины, в основном жители сельской местности. Основное требование к носителям языка — владение родным языком с детства и использование его в качестве основного средства общения.

Данные об информантах заносятся в таблицы со следующими полями:

- 1) порядковый номер;
- 2) фамилия, имя, отчество, или псевдоним;
- 3) код информанта;
- 4) пол;
- 5) возраст (год рождения, точный или приблизительный);
- 6) место рождения;
- 7) социальное происхождение;
- 8) образование;
- 9) профессия;
- 10) место постоянного проживания;
- 11) родовая принадлежность;
- 12) язык и диалект постоянного общения;
- 13) место записи;
- 14) год записи;

15) комментарии (информация относительно типа личности, речевых особенностей и т. п.);

16) информация о звуковых файлах, полученных от информанта.

Материал для записи состоит из двух блоков:

— чтение программы, составленной с учетом всех коммуникативных типов высказываний, различным актуальным членением и различной длины, порядка 60 предложений;

— разговорная речь — диалоги и монологи на повседневные темы. Примерные темы монологов: автобиография, семья и дети, праздники, сельскохозяйственные работы, исторические предания и др.

Аудиозапись осуществлялась при помощи цифровых диктофонов с использованием микрофонов. Параметры записи: 22 kHz, 16 bit, моно.

Информация об исходных звуковых файлах заносится в таблицы со следующими полями:

1) код информанта;

2) номер диктофонной записи (полученный при записи файла на диктофон);

3) общее время звучания;

4) общая информация о содержании (запись программы, рассказ о себе, диалоги на бытовые темы и т. п.);

5) комментарии к файлу;

6) наличие компьютерной обработки.

Полученные файлы подвергаются индексации — присвоению имен, в которых отражаются данные о диалекте, информанте, содержании и т. д., например: SrtM1-20043Int, где Srt означает сартульский диалект, M1 — код информанта, 20043 — номер исходного файла, Int — запись программы по интонации.

Сегментация и аннотирование

Аннотирование звуковых файлов производится в программе ELAN, которая широко используется в различных проектах по сохранению и документации языков [ELAN]. Эта программа позволяет создавать необходимое количество уровней аннотирования, а также может работать с интонационными графиками, созданными в программе PRAAT [PRAAT].

В Бурятском диалектном корпусе пока планируются следующие уровни аннотирования:

— макротема;

— текст в орфографической записи с сегментацией на синтагмы;

— иллокутивная функция синтагмы;

— текст в орфографической записи с сегментацией на слова;

— просодическая разметка (осуществляется в несколько уровней).

При первичном прослушивании звуковой файл делится на достаточно крупные тематические отрезки — тексты. Например, автобиография, рассказ о семье и детях, рассказ о языковой ситуации в данной местности, о возникновении названия населенных пунктов и т. д. Впоследствии будет составлен перечень всех макротем, полученных в результате тематического членения звучащего текста. Такая тематическая разметка позволит в будущем вести анализ характеристик звучащего диалектного текста.

Поскольку данный диалектный корпус задумывался нами прежде всего для исследования просодии, считаем целесообразным орфографическую расшифровку текста проводить на уровне синтагмы. Указываются следующие ее характеристики: завершенность (.), незавершенность (.), восклицательность (!), вопросительность (?) и т. д. Отмечаются паузы-хезитации, обрывы слов, растягивания звуков, неречевые явления (смех, кашель), переходы между языками (с бурятского на русский и обратно) и т. д.

Просодическая разметка

Учитывая направленность диалектологического звукового корпуса на исследование суперсегментных характеристик бурятской речи, в первоочередные задачи при его создании входит разработка принципов просодического аннотирования и системы просодической транскрипции. Просодическая транскрипция является способом «наглядно-слухового представления звучащей речи» (Брызгунова, 1984: 5), которая изначально использовалась в дидактических целях при обучении произношению, а в период активного создания звуковых корпусов и пристального внимания к особенностям звучащей речи стала необходима для просодического аннотирования корпусов.

Анализ просодических транскрипций, созданных ранее и применяемых в других звуковых корпусах, позволил разделить их на два типа:

— выбор из ограниченного числа интонационных моделей. Сюда можно отнести систему интонационных конструкций (ИК) Е. А. Брызгуновой (Брызгунова, 1977), впоследствии расширенную Н. Б. Вольской (Вольская, 2016). В таких транскрипциях предварительно устанавливается ограниченный набор наиболее типичных интонационных контуров, используемых в языке, с присвоением им определенного номера. Далее при просодическом аннотировании фраза членится на синтагмы, отмечается место интонационного центра и указывается номер интонационной модели;

— неопределенное количество интонаций, которые образуются путем сочетания элементарных просодических компонентов. Сюда может быть отнесена комбинаторная модель С. В. Кодзасова (Просодический строй русской речи: 124), а также Транскрипция русской интонации (ТоРИ) С. Оде (Оде, 1995). В таких транскрипциях основной задачей является выявление элементарных просодических компонентов и закономерностей их комбинирования.

В транскрипции С. В. Кодзасова, применяемой, к примеру, в Русском интонационном корпусе (RINCO), выделяются два типа просодий: интегральные — просодическая синтагма, и локальные — слоги, несущие тональные акценты. Для каждой из этих просодий указываются свои характеристики. Так, для интегральных просодий обозначаются: регистр, темп, громкость, типы фонаций, случаи редукции и т. п. Для локальных тональных акцентов отмечается направление движения тона на всех значимых акцентах во фразе, а также регистр, локализация, увеличенный или уменьшенный интервал и т. п. (Архипов, 2012).

Создатели современных корпусов разговорной речи при аннотировании материала почти всегда уделяют внимание просодическим характеристикам фразы. К примеру, проект «Рассказы о сновидениях и другие корпуса звучащей речи» хотя и не является чисто интонационным, а нацелен на описание и дискурсивное аннотирование устной речи, тем не менее содержит указания на важнейшие просодические характеристики речевого потока. В качестве минимального набора характеристик дискурсивной аннотации в работе принята сегментация речевого потока на элементарные дискурсивные единицы (ЭДЕ), и указание на ее иллокутивное значение: завершенность, незавершенность, восклицательность, вопрос, директива, обращение и т. д. При более подробной транскрипции к вышеуказанным добавляются важнейшие просодические характеристики: паузы и направление движения тона на словах, несущих главный акцент ЭДЕ, так как они непосредственным образом влияют на сегментацию речевого потока и определяют тип иллокутивного значения. Создатели транскрипции отмечают, что главные акценты выполняют три важные функции: выступают в качестве интонационного центра, поддерживающего целостность ЭДЕ и выделяют в нем наиболее значимую информацию, а направление движения тона определяет роль ЭДЕ в иллокутивной цепочке. В более полной транскрипции помимо движения тона на главном акценте добавляется характеристика движения тона на других словах, произносимых с акцентным выделением (Рассказы о сновидениях).

Анализ имеющихся транскрипций показал, что они отличаются друг от друга по способу представления материала, условными знаками и степенью подробности. Однако были выявлены основные способы и принципы транскрибирования, общие для большинства просодических разметок.

Прежде всего речевой поток должен быть сегментирован на речевые отрезки. В имеющихся транскрипциях эти отрезки называются по-разному: синтагмы, блоки, элементарные дискурсивные единицы (ЭДЕ) (Богданова-Бегларян, 2016). Для каждого отрезка обычно указывается место интонационного центра и участки просодической выделенности, а также описывается его мелодический рисунок. В некоторых просодических транскрипциях тип мелодического контура речевого отрезка выбирается из ограниченного набора интонационных контуров, ранее выявленных для данного конкретного языка, в других это могут быть комбинации элементарных просодических признаков, формирующих некоторое количество интонационных моделей.

Для просодической транскрипции, разрабатываемой для нужд Диалектного корпуса бурятского языка, были взяты за основу следующие принципы просодической транскрипции:

- членение фразы на синтагмы (завершенные, незавершенные, восклицательные, вопросительные и т. д.);
- выявление интонационного центра и участков просодической выделенности и описание мелодического контура на этих участках (/ — восходящий тон, \ — нисходящий тон, ^ — восходяще-нисходящий, v — нисходяще-восходящий, /- — восходяще-ровный, \- — нисходяще-ровный);
- указание на уровень тона на начале высказывания, интонационном центре и конце высказывания (H — высокий, M — средний, L — низкий);
- паузы заполненные (м-м, э-э) и абсолютные (А);
- ускорение или замедление темпа произнесения;
- повышение или понижение громкости.

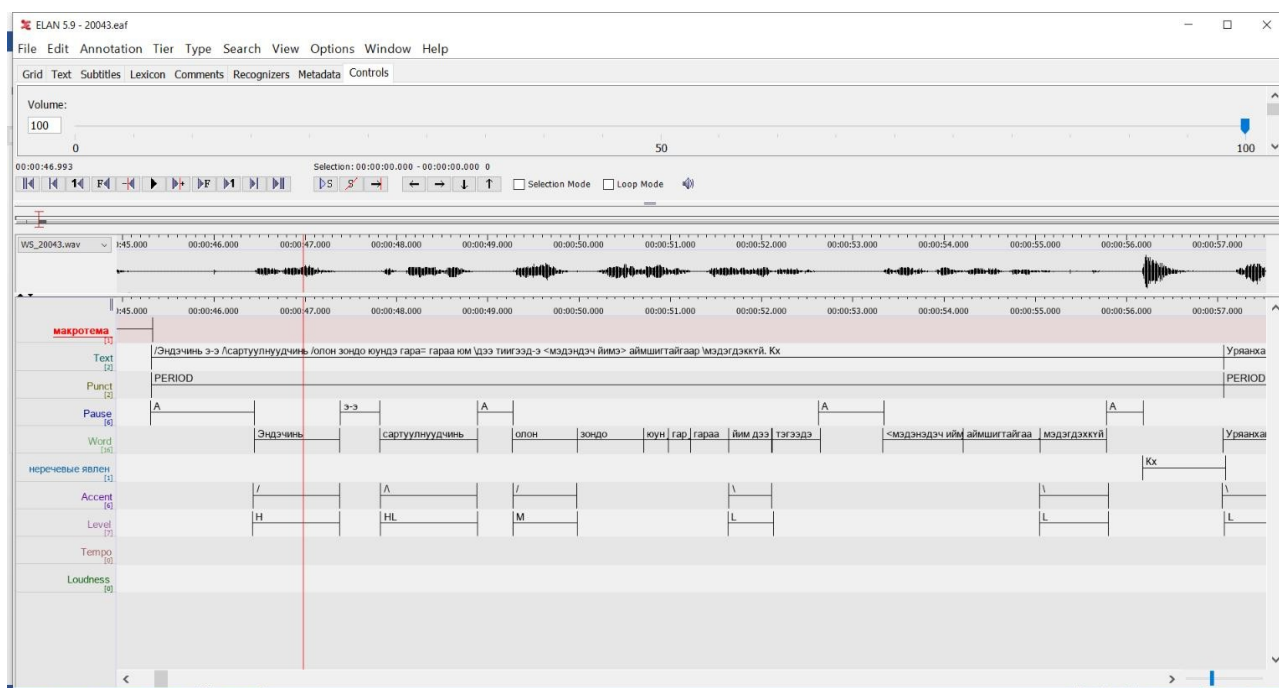


Рис. 1. Пример аннотации в программе ELAN

Заключение

Многодиалектный характер бурятского языка обусловил значительную вариативность бурятской разговорной речи, которая особенно заметна на лексическом и фонетическом уровнях. Создаваемый диалектный корпус послужит средством фиксации и изучения диалектной речи, что особенно актуально в условиях снижения уровня владения бурятским языком среди молодежи, а также отсутствием фундаментальных работ по бурятской диалектологии в последние десятилетия.

В настоящее время разработаны основные этапы работы над корпусом, принципы сбора звукового материала, уровни аннотирования. Учитывая направленность корпуса на изучение суперсегментных характеристик бурятской речи, особое внимание в корпусе уделяется принципам просодического аннотирования. Разработаны основные принципы системы просодической транскрипции для бурятского языка, которая будет в дальнейшем совершенствоваться.

Работа по созданию Звукового корпуса бурятских диалектов пока находится в начале пути. В дальнейшем уровни аннотирования будут корректироваться, материал корпуса расширяться, совершенствоваться методы исследования. Создание такого корпуса позволит расширить спектр изысканий в области бурятской диалектологии, сегментной и суперсегментной фонетики, зафиксировать современное состояние бурятских диалектов, проследить основные процессы, происходящие в речи современных бурят.

Список литературы

- Русский интонационный корпус: предварительный отчет [Текст] / А. В. Архипов, Л. М. Захаров, С. В. Кодзасов и др. // Компьютерная лингвистика и интеллектуальные технологии: матер. международ. конф. — М.: Изд-во РГГУ, 2012. — Вып. 11 (18). — С. 18–27.
- Бальжинимаева Ц. Ц.** Некоторые наблюдения над лексикой хоринского говора бурятского языка [Текст] / Ц. Ц. Бальжинимаева // Санжеевские чтения — 7: матер. международ. конф. — Улан-Удэ, 2012. — С. 18–25.
- Баторов Ц. В.** Об опыте создания первой газеты на диалектах бурятского языка [Текст] / Ц. В. Баторов // Язык и культура. — Новосибирск, 2016. — С. 7–20.
- Богданова-Бегларян Н. В.** Вопросительные конструкции в устной спонтанной речи: адресность, ритмичность, идиоматичность [Текст] / Н. В. Богданова-Бегларян // Коммуникативные исследования. 2016. № 4 (10). С. 61–76.
- Брызгунова Е. А.** Анализ русской диалектной интонации [Текст] / Е. А. Брызгунова // Экспериментально-фонетические исследования в области русской диалектологии. М.: Наука, 1977. 294 с.
- Будаев Ц. Б.** Бурятские диалекты (опыт диахронического исследования) [Текст] / Ц. Б. Будаев. — Новосибирск: Наука, 1992. — 217 с.
- Бураев И. Д.** Проблемы классификации бурятских диалектов [Текст] / И. Д. Бураев // Проблемы бурятской диалектологии. — Улан-Удэ: БНЦ СО РАН, 1996. — 190 с.
- Бурятский корпус [Электронный ресурс] — URL: http://web-corpora.net/BuryatCorpus/search/?interface_%20language=ru (Дата обращения 14.05.2019).
- Вольская Н. Б., Качковская Т. В.** Принципы просодической разметки в новом корпусе русской спонтанной речи SoRuSS [Текст] / Н. Б. Вольская // Фонетика сегодня. Материалы докладов и сообщений VIII международ. науч. конф. 2016. С. 29–31.
- Дырхеева Г. А.** Живая речь в Бурятии: к проблеме изучения (социолингвистический обзор) [Текст] / Г. А. Дырхеева // Вестник Томского государственного университета. Филология. — Томск: Национальный исследовательский Томский государственный университет, 2017. — № 45. — С. 77–88.
- Замбулаева Н. Г.** Тункинский говор в системе бурятских диалектов: морфологические особенности [Текст] / Н. Г. Замбулаева // Вестник Иркутского государственного лингвистического университета. — Иркутск, 2011. — № 3. — С. 69–76.
- Исследование бурятских говоров [Текст] / под ред. Ц. Б. Цыдендамбаева и И. Д. Бураева. — Улан-Удэ: Бурятское книжное издательство, 1965. — Вып. 1. — 189 с.

Исследование бурятских говоров [Текст] / под ред. Ц. Б. Цыдендамбаева, И. Д. Бураева — Улан-Удэ: Бурятское книжное издательство, 1968. — Вып. 2. — 196 с.

Оде С. Интонационная система русского языка в свете данных перцептивного анализа [Текст] / С. Оде // Проблемы Фонетики II. М., 1995. С. 200–215.

Патаева В. Д. Лексика народной медицины в тункинском говоре бурятского языка [Текст] / В. Д. Патаева // Вестник Бурятского государственного университета. — 2009. — № 10. — С. 26–30.

Просодический строй русской речи [Текст] / под ред. Т. М. Николаевой. М., 1996. 256 с.

Рассадин В. И. Очерки по исторической фонетике бурятского языка [Текст] / В. И. Рассадин. — М.: Наука, 1982. — 199 с.

Рассадин В. И. Присаянская группа бурятских говоров [Текст] / В. И. Рассадин. — Улан-Удэ, 1996. — 215 с.

Рассадин В. И. Становление говора нижеудинских бурят [Текст] / В. И. Рассадин. — Улан-Удэ: Изд-во БНЦ СО РАН, 1999. — 160 с.

Семенова В. И. О фонетических особенностях языка баяндаевских бурят [Текст] / В. И. Семенова // Национальная картина мира в литературе, фольклоре, языке и образовании: матер. регион. конф. — Иркутск, 2013. — С. 81–88.

Соктоева С. П. Консонантизм хоринского диалекта бурятского языка [Текст] / С. П. Соктоева. — Новосибирск: Наука, 1988. — 165 с.

Устные тексты на калмыцком языке: запись и расшифровка [Текст] / В. В. Куканова, Е. В. Бембеев, Н. Н. Убушаев и др. // Вестник Калмыцкого университета. — 2013. — № 3 (19). — С. 56–64.

Шерстинова Т. Ю., Люблинская М. Д. Звуковой корпус для изучения просодической системы младописьменных языков [Текст] / Т. Ю. Шерстинова, М. Д. Люблинская // Корпусная лингвистика — 2006: труды междунар. конф. — СПб., 2006. — С. 381–388.

ELAN (версия 5.9) [компьютерная программа]. (2020). Неймеген: Институт психолингвистики имени Макса Планка. URL: <https://archive.mpi.nl/tla/elan> (Дата обращения 28.05.2020).

Russian.news.cn [Электронный ресурс] — URL: http://russian.news.cn/2016-01/22/c_135036519.htm (Дата обращения 14.05.2019).

PRAAT [компьютерная программа]. URL: <http://www.fon.hum.uva.nl/praat/> (Дата обращения 28.05.2020).

Рассказы о свидениях и другие корпуса звучащей речи [Электронный ресурс]. URL: <http://spokencorpora.ru> (Дата обращения 10.09.2020).

The IViE Corpus [Электронный ресурс] — URL: <http://www.phon.ox.ac.uk/files/apps/IViE/> (Дата обращения 14.05.2019).

Абаева Юлия Догоржаповна, кандидат филологических наук, научный сотрудник.

Институт монголоведения, буддологии и тибетологии СО РАН.

Ул. Сахьяновой, д. 6, г. Улан-Удэ, Республика Бурятия, 670047.

E-mail: julaba@yandex.ru

Материал поступил в редакцию 19 июня 2019 г.

Iu. D. Abaeva

THE DIALECTAL CORPUS AS A MEAN OF MAINTAINING AND ANALYZING OF THE BURYAT PROSODY

The article is devoted to the creation of the Buryat dialectal corpus. This Corpus is mainly oriented on the study of the prosody of the Buryat dialects. Although the dialects are actively functioning, their current state is a matter of concern. However, in the 80-s of the last century, the program for studying dialects was reduced, and till nowadays there are no large-scale, systematic researches in the field. The creation of such resource is due to the getting worse sociolinguistic situation with the Buryat dialects, as well as the lack of modern researches of the Buryat prosody and in the Buryat dialectology in general.

The paper describes the main stages of work on the Corpus, requirements for the collection of material, the principles of its segmentation and annotation. The material was recorded in areas of Buryat territory, speakers both men and women, who know their dialect since childhood and use it in everyday communication. The recording program consists of two blocks: prepared phrases for the study of

intonation and spontaneous everyday speech. The intonation analysis is carried out using the PRAAT program. For segmentation and annotation the program ELAN is used.

The work pays special attention to the principles of prosodic markup of Buryat. Available prosodic transcriptions were analyzed and basic principles of prosodic annotation were developed. In the Buryat dialectal corpus prosodic transcription will include: segmentation of the speech into phrases which differ in their illocutionary function, indication of the location of the intonation center and areas of prosodic emphasis, the description of the pitch contour. Special signs were developed.

The work on the creation of the Dialect Corpus is at the beginning, so plans and future prospects are outlined. This resource will be useful not only for the study of the dialect intonation, but also for the Buryat dialectology in general.

Keywords: *Buryat dialects, spoken dialectal corpus, prosody, stress, intonation, prosodic annotation.*

References

- Arhipov A. V., Zakharov L. M., Kodzasov S. V., Lebedev A. A.** (2012). Russkiy intonacionniy korpus: predvaritel'niy otchet [Russian Intonation Corpus: a Preliminary Report]. *Komp'yuternaya lingvistika i intellektual'nye tehnologii: mater. mezhdunarod. konf.* [Proc. of the International Conference in Computational linguistics and intellectual technologies «Dialogue»] — M.: Izd-vo RGGU, 2012. — Vyp. 11 (18). — S. 18–27.
- Bal'zhinimaeva Ts. Ts.** (2012). Nekotorye nabljudeniya nad leksikoy horinskogo govora burjatskogo yazyka [Some observations on the vocabulary of the Khorinsky dialect of the Buryat language]. *Sanzheevskie chteniya — 7: mater. mezhdunar. konf.* [Proc. of the International Conference «Sanzheev's readings — 7»] — Ulan-Ude, 2012. — S. 18–25.
- Batorov Ts. V.** (2016). Ob opyte sozdaniya pervoy gazety na dialektakh buryatskogo yazyka [On the experience of creating of the first newspaper in dialects of the Buryat language]. *Yazyk i kul'tura* [Language and Culture]. — Novosibirsk, 2016. — S. 7–20.
- Bogdanova-Beglaryan N. V.** (2016). Voprositel'nye konstrukcii v ustnoj spontannoj rechi: adresnost', ritmichnost', idiomatichnost' [Interrogative structures in oral spontaneous speech: targeting, rhythm, idiomatic]. *Kommunikativnye issledovaniya* [Communicative research]. 2016. № 4 (10). S. 61–76.
- Bryzgunova E. A.** (1977). Analiz russkoj dialektnoj intonacii [Analysis of Russian dialect intonation]. *Ehksperimental'no-foneticheskie issledovaniya v oblasti russkoj dialektologii.* [Experimental and phonetic research in the field of Russian dialectology] M.: Nauka, 1977. 294 s.
- Budaev Ts. B.** (1992). Buryatskie dialekty (opyt diakhronicheskogo issledovaniya) [Buryat dialects (experience of diachronic research)]. — Novosibirsk: Nauka, 1992. — 217 s.
- Buraev I. D.** (1996). Problemy klassifikacii burjatskih dialektov [Problems of classification of the Buryat dialects]. *Problemy burjatskoj dialektologii* [Problems of the Buryat dialectology]. — Ulan-Ude: BNC SO RAN, 1996. — 190 s.
- Burjatskiy korpus [The Buryat Corpora]. Retrieved May 14, 2019 from: http://web-corpora.net/BuryatCorpus/search/?interface_%20language=ru.
- Vol'skaya N. B., Kachkovskaya T. V.** Printsipy prosodicheskoy razmetki v novom korpuse russkoy spontannoj rechi CoRuSS [The principles of prosodic annotation in a new corpus of Russian spontaneous speech]. *Fonetika segodnya.* [Phonetics today]. Materialy dokladov i soobshcheniy VIII mezhdunarod. nauch. konf. 2016. S. 29–31.
- Dyrkheeva G. A.** (2017). Zhivaya rech' v Burjatii: k probleme izucheniya (sociolingvisticheskiy obzor) [Spoken speech in Buryatia: to the problem of analyzing (sociolinguistic review)]. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya* [Tomsk State University Journal of Philology] — Tomsk: Nacional'nyj issledovatel'skiy Tomskij gosudarstvennyj universitet, 2017. — № 45. — S. 77–88.
- ELAN (version 5.9) [computer program]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics. URL: <https://archive.mpi.nl/tla/elan> (Date of apply 05/28/2020)
- Issledovanie burjatskih govorov [Study of the Buryat dialects] / pod red. Ts. B. Tsydendambaeva i I. D. Buraeva. — Ulan-Ude: Burjatskoe knizhnoe izdatel'stvo, 1965. — Vyp. 1. — 189 s.
- Issledovanie burjatskih govorov [Study of the Buryat dialects] / pod red. Ts. B. Tsydendambaeva, I. D. Buraeva — Ulan-Ude: Burjatskoe knizhnoe izdatel'stvo, 1968. — Vyp. 2. — 196 s.
- Kukanova V. V., Bembeev E. V., Ubushaev N. N., Mandzhiyeva B. B.** (2013). Ustnye teksty na kalmyskom yazyke: zapis' i rasshifrovka [Oral texts in the Kalmyk language: recording and decoding]. *Vestnik Kalmytskogo universiteta* [Bulletin of Kalmyk State University]. — 2013. — № 3 (19). — S. 56–64.

- Ode C.** Intonatsionnaya sistema russkogo yazyka v svete dannykh pertseptivnogo analiza [The intonation system of the Russian language in the light of the data of perceptual analysis] // *Problemy Fonetiki II*. [Issues in Phonetics II] М., 1995. S. 200–215.
- Pataeva V. D.** (2009). Leksika narodnoy mediciny v tunkinskom govore buryatskogo yazyka [Vocabulary of traditional medicine in Tunkinskii dialect of the Buryat language]. *Vestnik Buryatskogo gosudarstvennogo universiteta* [Bulletin of the Buryat State University]. — 2009. — № 10. — S. 26–30.
- PRAAT [computer program]. URL: <http://www.fon.hum.uva.nl/praat/> (Date of treatment 05/28/2020).
- Prosodicheskiy stroy russkoy rechi [Prosodic structure of Russian speech] / pod red. T. M. Nikolayevoy. М., 1996. 256 s.
- Rassadin V. I.** (1982). Ocherki po istoricheskoy fonetike buryatskogo yazyka [Essays on the historical phonetics of the Buryat language]. — М.: Nauka, 1982. — 199 s.
- Rassadin V. I.** (1996). Prisyanskaya gruppa buryatskikh govorov [Sayan group of the Buryat dialects]. — Ulan-Ude, 1996. — 215 s.
- Rassadin V. I.** (1999). Stanovlenie govora nizhneudinskikh buryat [Formation of the Nizhneudinsk Buryat dialect]. — Ulan-Ude: Izd-vo BNC SO RAN, 1999. — 160 s.
- Rasskazy o snovideniyakh i drugie korpusa zvuchashchey rechi [Dream stories and other corpuses of spoken speech]. URL: <http://spokencorpora.ru> (Data of applying 10.09.2020).
- Russian.news.cn. Retrieved May 14, 2019 from: http://russian.news.cn/2016-01/22/c_135036519.htm.
- Semenova V. I.** (2013). O foneticheskikh osobennostyakh yazyka bayandaevskikh buryat [On the phonetic features of the language of the Bayanday Buryats]. *Nacional'naya kartina mira v literature, fol'klore, yazyke i obrazovanii: mater. region. konf.* [Proc. of the Regional Conference National world view in literature, folklore, language and education]. — Irkutsk, 2013. — S. 81–88.
- Sherstinova T. Ju., Ljublinskaja M. D.** (2006). Zvukovoy korpus dlya izucheniya prosodicheskoy sistemy mladopis'mennykh yazykov [Sound corpus for studying of the prosodic system of young-written languages]. *Korpusnaya lingvistika — 2006: trudy mezhdunarod. konf.* [Proc. of the International Conference Corpus Linguistics — 2006] — SPb., 2006. — S. 381–388.
- Soktoeva S. P.** (1988). Konsonantizm khorinskogo dialekta buryatskogo yazyka [Consonant system of the Khorinsky dialect of the Buryat language]. — Novosibirsk: Nauka, 1988. — 165 s.
- The IViE Corpus. Retrieved May 14, 2019 from: <http://www.phon.ox.ac.uk/files/apps/IViE/>.
- Zambulaeva N. G.** (2011). Tunkinskiy govor v sisteme buryatskikh dialektov: morfologicheskie osobennosti [Tunkinsky dialect in the system of Buryat dialects: morphological features]. *Vestnik Irkutskogo gosudarstvennogo lingvisticheskogo universiteta*. [Vestnik of Irkutsk State Linguistic University] — Irkutsk, 2011. — № 3. — S. 69–76.

Abaeva Iuliia Dogorzhapovna, candidate of philology, research fellow.

RAS, Siberian branch, Institute for Mongolian, Buddhist and Tibetan Studies.

6 Sakhyanovoy st., Ulan-Ude, Russia, 670047.

E-mail: julaba@yandex.ru