

<https://doi.org/10.30853/filnauki.2019.10.35>

Абаева Юлия Догоржаповна

ТЕРМИНЫ КОРПУСНОЙ ЛИНГВИСТИКИ В ХАЛХА-МОНГОЛЬСКОМ И БУРЯТСКОМ ЯЗЫКАХ

В статье проводится сравнительный анализ терминов корпусной лингвистики в халха-монгольском и бурятском языках. Приводится краткий обзор основных монгольязычных корпусов в России, Монголии и КНР. Исследование терминов халха-монгольского языка показало недостаточную разработанность этой лексики: параллельное употребление заимствованных и монгольских слов, наличие нескольких синонимических вариантов одного термина. Аналогичные термины в бурятском языке только начинают разрабатываться. Сравнение их с халха-монгольскими терминами выявило возможности создания новых терминов, потенциальные заимствования из общемонгольской лексической терминологии.

Адрес статьи: www.gramota.net/materials/2/2019/10/35.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2019. Том 12. Выпуск 10. С. 162-165. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2019/10/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net
Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: phil@gramota.net

Языки народов Российской Федерации

The Russian Federation Peoples' Languages

УДК 811.512.3

Дата поступления рукописи: 29.08.2019

<https://doi.org/10.30853/filnauki.2019.10.35>

В статье проводится сравнительный анализ терминов корпусной лингвистики в халха-монгольском и бурятском языках. Приводится краткий обзор основных монголоязычных корпусов в России, Монголии и КНР. Исследование терминов халха-монгольского языка показало недостаточную разработанность этой лексики: параллельное употребление заимствованных и монгольских слов, наличие нескольких синонимических вариантов одного термина. Аналогичные термины в бурятском языке только начинают разрабатываться. Сравнение их с халха-монгольскими терминами выявило возможности создания новых терминов, потенциальные заимствования из общемонгольской лексической терминологии.

Ключевые слова и фразы: корпусная лингвистика; корпус; база данных; термины; терминологизация; калькирование; заимствование; халха-монгольский язык; бурятский язык.

Абаева Юлия Догоржаповна, к. филол. н.

Институт монголоведения, буддологии и тибетологии Сибирского отделения Российской академии наук,

г. Улан-Удэ

julaba@yandex.ru

ТЕРМИНЫ КОРПУСНОЙ ЛИНГВИСТИКИ В ХАЛХА-МОНГОЛЬСКОМ И БУРЯТСКОМ ЯЗЫКАХ

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

В монгольском языкознании активно ведется работа по созданию корпусов. В России это Корпус монгольского языка (создаваемый С. А. Крыловым) [9], Бурятский корпус (составитель Л. Д. Бадамаева) [6], Калмыцкий корпус (составитель А. Э. Ванькаева) [7], которые располагаются на одном сервере <http://web-corpora.net/> и используют одну поисковую платформу [1, с. 150]. Национальный корпус калмыцкого языка – еще один калмыцкий корпус (основной разработчик В. В. Куканова) [10], который располагается на самостоятельном сервере и имеет свою поисковую систему [8]. По оценке Л. Д. Бадмаевой, на сегодняшний день он является наиболее разработанным среди российских монголоязычных корпусов [1, с. 150].

В Монголии в Центре компьютерной лингвистики Монгольского госуниверситета был создан корпус, состоящий из 5 млн слов, куда вошли тексты художественной литературы, законы, публицистика, газетные публикации. В Академии наук Монголии ведется работа над Корпусом монгольского языка под руководством Т. Пурэвсүрэн на материале избранных сочинений современной монгольской литературы, состоящих из 108 томов [11].

В Китайском монголоведении корпусные лингвистические исследования начались значительно раньше и, нужно отметить, продвинулись намного дальше, чем в других монголоведных центрах. Так, ученые Университета Внутренней Монголии с 1984 года начали создавать Базу данных современного монгольского языка. На сегодняшний день в результате 30-летней непрерывной работы эта база насчитывает 10 млн слов. Среди ученых, внесших вклад в развитие корпусной лингвистики в Китае, нужно отметить проф. Чойжинжав, проф. С. Насан-Урт, проф. Хуашабу и др.

Институтом изучения монгольского языка Академии социальных наук Автономного района Внутренняя Монголия КНР в г. Хух-Хото в 2005 году начата разработка Корпуса монгольских языков, цель которого – сделать синхронный срез современного состояния монгольских языков, существующих в мире. Группу ученых возглавляет лексикограф Д. Баатар. В отличие от вышеуказанного корпуса, этот корпус является звуковым, то есть ориентирован на сохранение звучащей речи во всем ее многообразии. Материал записывался в провинциях и автономных районах Китая, аймаках Монголии, районах Республик Бурятия и Калмыкия.

Актуальность настоящего исследования обусловлена активным развитием корпусных лингвистических исследований, в связи с чем назрела задача создания новых терминов, анализа уже сложившейся

терминологической лексики и унификации ее в монгольских языках. Перед монголоведами разных стран остро стоит вопрос о стандартизации терминологического аппарата, что является залогом успешной и продуктивной совместной научной деятельности, позволит вести диалог без искажения смысла. Функционирование монгольских языков в разных экстралингвистических условиях не способствует созданию единых и однозначных терминов, а тот факт, что корпусная лингвистика, как и компьютерная лингвистика вообще, является новой областью знания, терминология которой еще до конца не разработана, еще более усугубляет ситуацию с терминологическими разночтениями, синонимией и дублетами.

Цель статьи – сравнительный анализ терминологической лексики, связанной с корпусной лингвистикой, в халха-монгольском и бурятском языках на материале научных и научно-публицистических текстов. **Задачи** настоящего исследования заключаются в анализе способов образования терминов корпусной лингвистики, а также сравнении уже устоявшихся терминов, употребляющихся при характеристике корпусов и этапов работы над ними; внесении предложений по унификации терминов, расширению терминологической базы бурятского языка.

Подобное исследование терминов корпусной лингвистики в бурятском языке проводится впервые. По халха-монгольскому языку есть статья Д. Боролзой [3], посвященная компьютерной лингвистике и вопросам создания электронных корпусов, в которой имеются некоторые замечания о вариантах передачи понятий «корпус», «база данных» на халха-монгольском языке. Нужно отметить, что бурятскоязычного материала по корпусной лингвистике, как и научных текстов вообще, не так много. Вследствие этого корпусная терминология в бурятском языке практически не развита. Такой сравнительный анализ будет полезен именно для бурятского языка, позволит ввести в научный оборот новую лексику, уточнит смысл имеющихся терминов, снизит количество синонимов и дублетов.

Один из вариантов толкования английского слова *corpus* ‘корпус’ – «собрание письменных или устных материалов, хранимых в компьютере и используемых для изучения закономерностей функционирования языка» (перевод автора статьи. – Ю. А.) [13]. Этот термин в монгольских языках активно употребляется в виде заимствования: бур. *Буряад хэлэнэй корпус* ‘Корпус бурятского языка’; бур. *корпус хэлэ шэнжэлэлэ*, х.-монг. *корпус хэл шинжлэл* ‘корпусная лингвистика’: х.-монг. *1980-аад оноос корпус хэл шинжлэл нь ихээхэн хурдацтай хөгжижэ* [3]... / «С 1980-х гг. корпусная лингвистика очень быстро развивалась...».

Однако в халха-монгольском языке более активно используется монгольский вариант этого термина, образованный путем терминологизации слова *хөмрөг* с первоначальным значением ‘кладовая (для зерна), амбар, житница, закрома; казначейство; фонд’ [2, с. 130]: *БНХАУ-ын монголчуудын хувьд хөмрөг байгуулалтын ажил нь нэлээд эрт эхэлж...* / «Что касается монголов КНР, то работа по созданию корпуса у них началась намного раньше...». Термин *хөмрөгийн хэл шинжлэл* ‘компьютерная лингвистика’ в халха-монгольском языке практически вытеснил вариант этого термина с заимствованным словом «корпус»: *Компьютер хэл шинжлэл, тэр дундаа хөмрөгийн хэл шинжлэлийн судалгаа харьцангуй өндөр түвшинд хийгдэж буй* [11, х. 107]. / «Компьютерная лингвистика, в частности исследования корпусной лингвистики, находится на относительно высоком уровне». В бурятском языке используется только вариант с заимствованным словом: *корпус хэлэ шэнжэлэлэ*.

Сан ‘казна, фонд, амбар’ – еще одно монгольское слово, эквивалентное слову «корпус». Как пишет Г. Ц. Пюрбеев, «в закреплении терминологического значения слова решающую роль играет контекст, который помогает установить принадлежность термина к той или иной области профессиональной лексики» [12, х. 338]. Для создания соответствующего контекста слово *сан* в значении «корпус» обычно употребляется с определениями *материалын сан, хэлний материалын сан* ‘догл. хранилище языкового материала’. Например: *Хятад улсын хувьд хэлний материалын сан байгуулах ажил 1980-аад оноос эхэлсэн байна* [3]. / «Что касается Китая, то работа по созданию языкового корпуса началась в 1980-х гг.». Эти же определения используются и со словом *хөмрөг* в значении «корпус»: *хэлний хөмрөг, хэлний материалын хөмрөг*. Например: *Хэлний хөмрөгийг ашиглан хэлний материалд суурилж хэлийг судалдаг ухааныг корпус хэл шинжлэл гэдэг* [Там же]. / «Науку, изучающую язык с использованием языкового материала из языкового корпуса, называют корпусная лингвистика».

Слово «корпус» часто употребляется с определением: х.-монг. *цахим*, бур. *сахим* ‘электронный’: бур. *Мүнөө сагай шэнэ шэнжэлэлгын ажлал гэхэдэ, буряад хэлэнэй сахим корпус болохо юм*. / «Если говорить о современных научных разработках, нужно упомянуть электронный корпус бурятского языка»; х.-монг. *Бидний байгуулжэ буй монгол хэлний цахим хөмрөг нь орчин цагийн монгол утга зохиолын хэлний гол бүрэлдүүлэгч болох уран зохиолын материалд түшиглэсэн, цуваа цагийн, тусгай хөмрөг юм*. / «Составляемый нами электронный корпус монгольского языка представляет собой специальный диахронический корпус, основанный на литературных материалах, которые являются основными составляющими современной монгольской литературы» [11, х. 109]; х.-монг. *Компьютер хэл шинжлэл ба цахим сан байгуулах асуудалд* [3]. / «К вопросу о компьютерной лингвистике и создании электронного корпуса».

Основное понятие корпусной лингвистики «база данных» (англ. *database*) в монгольском языке часто используется без перевода как *дата бааз*: х.-монг. *АНУ-аас гадна байдаг анхны дата бааз*. / «Первая база данных за пределами США». Однако это словосочетание имеет следующие варианты на халха-монгольском опять же с использованием слов *хөмрөг* и *сан*:

дата хөмрөг – слово «база» заменяется на *хөмрөг* ‘фонд’: *Одоо үеийн монгол хэлний дата хөмрөгийн программ*. / «Программа для создания базы данных современного монгольского языка»;

мэдээллийн сан или *хэлний мэдээллийн сан*, где *мэдээлэл* ‘информация’, *мэдээллийн* ‘информационный’: *...database-ийг мэдээллийн сан хэмээн орчуулсан байна*. / «...database перевели как “мэдээллийн сан”» [Там же];

мэдээллийн хөмрөг: х.-монг. *Материалын сан гэдэг нь хэлний материалыг хадагалах сан буюу мэдээллийн хөмрөг юм.* / «Корпус представляет собой хранилище языкового материала или базу данных»;

өгөгдөлийн сан, где *өгөгдөл* ‘данные’: *Гэвч «машин сурах» гэдэг нь зөвхөн өгөгдөлийн сантай холбоотой асуудал биш бөгөөд энэ нь бас хиймэл оюун ухаантай холбогдоно* [11, х. 110]. / «Однако вопрос так называемого “обучения машин” связан не только с базами данных, но также и с искусственным интеллектом».

Как видим, в халха-монгольском языке слова *сан* и *хөмрөг* употребляются и в значении «база данных», и в значении «корпус». В значении «корпус» к ним обычно присоединяются определения *хэлний* ‘языковой’, *материалын* ‘материалов’, *хэлний материалын* ‘языковых материалов’. При образовании словосочетания «база данных» используются слова *мэдээллийн*, *өгөгдөлийн*, а также заимствование *дата* (англ. *data*). Вариант *цахим сан* представляет собой термин описательного характера, передающий суть корпусов или баз данных как электронного (компьютерного) хранилища.

Слова *сан* и *хөмрөг* в халха-монгольском синонимичны и взаимозаменяемы: *Дээрх шаардлагыг хангасан өгөгдөлийн сан буюу хөмрөг* – *ийг үүсгэхдээ...* / «При создании “өгөгдөлийн сан буюу хөмрөг” (базы данных или хранилища), отвечающей вышеуказанным требованиям...» [Там же]; *Халимаг хэлний үндэсний хөмрөг*. / «Национальный корпус калмыцкого языка» [Там же, х. 109]; *Британы үндэсний сан*. / «Британский национальный корпус» [3, с. 3]. Однако в исследованиях последних лет прослеживается тенденция дифференциации значений этих слов. К примеру, Т. Пурэвсурэн, описывая структуру создаваемого ею корпуса, ранжирует слова, обозначая общую работу словом *хөмрөг*, а более мелкие ответвления – словом *сан*. Например, *Жишиг хөмрөг* ‘Корпус-образец’ состоит из *үгсийн сан* ‘лексический подкорпус’ и *хэлзүйн сан* ‘грамматический подкорпус’ [11, х. 109].

В бурятском языке следующие слова могут быть использованы при создании терминов «корпус» и «база данных»: *сан* и его фонетический вариант *һан* ‘хранилище, фонд’, *жаса* ‘фонд, собрание’, *нөөсэ* ‘накопления, запас, фонд’. Слова, соответствующего монгольскому *хөмрөг*, нет. Хотя обычно понятие «корпус» в бурятском языке передается заимствованным словом, можно предложить вариант по аналогии с халха-монгольскими терминами: *хэлэнэй материалал сахим сан* ‘электронное хранилище языковых материалов’, или более кратко *материалал сан* ‘хранилище материала’, или *хэлэнэй сан* ‘хранилище языка’. Кроме того, предполагается целесообразным внедрение в бурятский язык заимствования *хөмрөг* в виде *хүмэргэ* в значении «корпус».

Слово *database* ‘база данных’ в бурятском языке в виде иностранного заимствования *дата бааза*, аналогичного халха-монгольскому, не употребляется. Дословный перевод словосочетания «база данных» на бурятский язык выглядит следующим образом: «данные» – *баримта* или *мэдээ баримта*; «база» в значении «место хранения» не переводится, обозначается заимствованием *бааза*. Соответственно, термин «база данных» на бурятский можно перевести как *баримтын бааза* или *мэдээ баримтын бааза*. Если же взять за образец халха-монгольский вариант *мэдээллийн сан*, то в бурятском языке возможен вариант *мэдээнэй сан* или *мэдээнэй сан нөөсэ*.

По наполнению и способу представления материала корпусы могут быть:

устными (х.-монг. *яриа хэлний*, бур. *ярянай хэлэнэй*, *аман угын*, *аман хэлэлгын*) и письменными (х.-монг. *бичгийн*, бур. *бэшэгэй*);

фольклорными (х.-монг. *ардын аман зохиолой*, бур. *аман зохёолой*) и литературными (х.-монг. *утга зохиолой*, *уран зохиолой*, бур. *уран зохёолой*);

синхроническими (х.-монг. *хам цагийн*, бур. *синхронно*) и диахроническими (х.-монг. *цуваа цагийн*, бур. *диахронно*);

диалектными (х.-монг. *нутгийн аялгууны*, бур. *нутагай хэлэнэй* либо *нутагай аман аялгын*) и литературными (х.-монг. *бичгийн хэлэнэй*, бур. *бэшэгэй хэлэнэй*).

Корпус – это не просто собрание текстов, материал корпуса должен быть соответствующим образом аннотирован или размечен. В халха-монгольском языке в значении «аннотировать» употребляется *боловсруулах* ‘обрабатывать’: *Корпус хэл шинжлэл нь хэлний материалын хөмрөгийг байгуулах, хэлний материалыг боловсруулах, хэлний материалыг ашиглах зэргийг судалдаг* [3]. / «Корпусная лингвистика изучает создание корпуса языкового материала, аннотирование языкового материала, использование языкового материала и т.п.». Возможные бурятские варианты этого термина – *болбосоруулха*, *тэмдэглэхэ*, *аннотаци хэхэ*.

Уровни аннотирования корпуса (х.-монг. *үе шата*, бур. *үе шата*, *хуби һалбаринууд*):

– буквы/звуки (х.-монг. *үсэг* / *авиа боловсруулалт*, бур. *үзэг* / *абья болбосоруулга* / *тэмдэглэлэгэ*);

– слово (х.-монг. *үг боловсруулалт*, бур. *үгэ болбосоруулга* / *тэмдэглэлэгэ*);

– предложение (х.-монг. *өгүүлбэр боловсруулалт*, бур. *мэдүүлэл болбосоруулга/тэмдэглэлэгэ*);

– текст (х.-монг. *цогцолбор* или *эх боловсруулалт*, бур. *текст болбосоруулга* / *тэмдэглэлэгэ*).

Например: *Ажлын явц нь үсэг боловсруулалт, үг боловсруулалт, өгүүлбэр боловсруулалт, эх боловсруулалт гэсэн үе шаттай байна* [Там же]. / «Работа состоит из нескольких уровней: аннотирование букв, аннотирование слов, аннотирование предложений, аннотирование текста».

Как видно из приведенных примеров, для обозначения понятий, связанных с видами корпусов и этапов работы над ним, используются уже устоявшиеся термины бурятского и халха-монгольского языков, которые в целом идентичны, однако имеются некоторые различия в обозначении одного и того же понятия, например: х.-монг. *өгүүлбэр боловсруулалт* / бур. *мэдүүлэл болбосоруулга*. Для бурятского языка характерно большее количество синонимических вариантов: ср. *ярянай хэлэнэй*, *аман угын*, *аман хэлэлгын* ‘устный’; бур. *нутагай хэлэнэй*, *нутагай аман аялгын* ‘диалектный’.

Следует отметить, что в бурятском языке не разработаны термины «синхронический» и «диахронический», аналогичные халха-монгольским *хам цагийн* и *цүваа цагийн*. В х.-монг. *хам* 'единый', *цүваа* 'непрерывно идущий, тянущийся'. В бурятском языке слово *хам* в подобном значении не употребляется, слово *субаан* имеет значение 'вереница, непрерывное хождение', вероятно, возможно заимствование этого термина в виде *субаа саг* 'диахрония'. В бурятском языке эти термины употребляются в виде заимствования *диахронно* 'диахронный', *синхронно* 'синхронный' либо носят скорее вид толкования: *удаан сагта хубилан хугжэлгэ* – дословно 'развитие в течение длительного времени' [4, с. 275]; *тон нэгэ сагта / нэгэн зэргэ болодог*, что дословно означает 'происходящее одновременно' [5, с. 170].

Таким образом, анализ терминов, связанных с новой областью лингвистики – корпусными исследованиями, показал, что в халха-монгольском языке эти термины находятся в стадии становления, о чем свидетельствует параллельное употребление заимствованных и монгольских слов (*корпус* / *хөмрөг*, *дата бааз* / *өгөгдөлийн сан*), синонимическое употребление слов (*хөмрөг* – *сан*) и наличие нескольких синонимических вариантов одного термина (*дата бааз*, *дата хөмрөг*, *мэдээллийн сан*, *өгөгдөлийн сан*). Выявлены следующие способы образования терминов: заимствования – х.-монг. *корпус*, *датабааз*, бур. *корпус*, *синхронно*, *диахронно*; терминологизация общеупотребительных слов – х.-монг. *хөмрөг*, х.-монг., бур. *сан*, х.-монг. *боловсруулах*; калькирование – х.-монг. *мэдээллийн / өгөгдөлийн сан*. В бурятском языке подобная терминология развита слабо, поэтому сравнительный ее анализ с терминами халха-монгольского языка способствует внедрению новой лексики. Думается, для бурятского языка было бы целесообразно заимствовать из монгольских языков термин *хөмрөг*, вероятно, в виде *хүмэргэ*, а также термины «синхрония» *субаа саг* и «диахрония» *хам саг*. Разработка единых принципов терминологии, унификация имеющихся терминов позволит языковедам из России и Монголии вести совместные исследования, обмениваться научными результатами и профессиональной информацией.

Список источников

1. Бадмаева Л. Д. Монголызычные корпуса: современное состояние // Вестник Бурятского государственного университета. 2015. № 10. С. 148-152.
2. Большой академический монгольско-русский словарь: в 4-х т. / отв. ред. Г. Ц. Пюрбеев. М.: Академия, 2002. Т. 4. 502 с.
3. Боролзой Д. Компьютер Хэл Шинжлэл Ба Цахим Үгийн Сан Байгуулах Асуудалд [Электронный ресурс]. URL: [https://borolzoi.blogmn.net/27067/kompiyutyer-hel-shinjllel-ba-tzahim-ugiin-san-\(1\)-baiguulah-asudald.html](https://borolzoi.blogmn.net/27067/kompiyutyer-hel-shinjllel-ba-tzahim-ugiin-san-(1)-baiguulah-asudald.html) (дата обращения: 10.07.2019).
4. Буряад-оруд толи. Бурятско-русский словарь: в 2-х т. Улан-Удэ: Республиканская типография, 2010. Т. 1. А – Н. 636 с.
5. Буряад-оруд толи. Бурятско-русский словарь: в 2-х т. Улан-Удэ: Республиканская типография, 2010. Т. 2. О – Я. 708 с.
6. Бурятский корпус [Электронный ресурс]. URL: http://web-corpora.net/BuryatCorpus/search/?interface_%20language=ru (дата обращения: 11.07.2019).
7. Калмыцкий корпус [Электронный ресурс]. URL: http://web-corpora.net/KalmykCorpus/search/?interface_language=ru (дата обращения: 20.09.2019).
8. Куканова В. В., Бембеев Е. В., Мулаева Н. М., Очирова Н. Ч. Национальный корпус калмыцкого языка: архитектура и возможности использования // Вестник Калмыцкого института гуманитарных исследований РАН. 2012. № 3. С. 138-150.
9. Монгольский корпус [Электронный ресурс]. URL: http://web-corpora.net/MongolianCorpus/search/?interface_language=ru (дата обращения: 11.07.2019).
10. Национальный корпус калмыцкого языка [Электронный ресурс]. URL: <http://kalmcorp.ru/> (дата обращения: 11.07.2019).
11. Пүрэвсүрэн Т. Монгол хэлний цахим хөмрөгийн байгуулалт // Хэл зохиол судлал: хэл шинжлэл, утга зохиол судлалын эрдэм шинжилгээний цуврал / еронхий ред. Г. Билгуудэй. Улаанбаатар: Соёмбо Принтинг, 2018. Х. 106-125.
12. Пюрбеев Г. Ц. История и современное состояние калмыцкой терминологии // Монгол-Орос хэл шинжлэлийн түүвэр / ред. Г. Билгуудэй и др. Улаанбаатар: Соёмбо принтинг, 2018. Х. 332-343.
13. Corpus [Электронный ресурс] // Cambridge Dictionary. URL: <https://dictionary.cambridge.org/ru/словарь/английский/corpus> (дата обращения: 11.07.2019).

CORPUS LINGUISTICS TERMS IN THE KHALKHA MONGOLIAN AND BURYAT LANGUAGES

Abaeva Yuliya Dogorzhapovna, Ph. D. in Philology

*Institute of Mongolian, Buddhist and Tibetan Studies of the Siberian Branch of the Russian Academy of Sciences, Ulan-Ude
julaba@yandex.ru*

The article provides a comparative analysis of corpus linguistics terms in the Khalkha Mongolian and Buryat languages. The author presents a brief survey of the basic Mongolian corpora in Russia, Mongolia and the People's Republic of China. The analysis of the Khalkha Mongolian terms has indicated insufficient assimilation of these lexical units: parallel usage of the Mongolian and borrowed words, existence of several synonymic variants of one and the same term. Analogous terms in the Buryat language are just being formed. The comparative analysis of the Buryat and Khalkha Mongolian terms has allowed the author to reveal a term-formative potential, to identify possible borrowings from the all-Mongolian terminology.

Key words and phrases: corpus linguistics; corpus; database; terms; terminologization; calque; borrowing; Khalkha Mongolian language; Buryat language.