

Современные компьютерные возможности в изучении башкирского языка и культуры

Борис Валерьевич Орехов

кандидат филологических наук

Национальный исследовательский университет «Высшая
школа экономики»

nevmenandr@gmail.com

План разговора:

1. Корпуса башкирского языка:

- Национальный корпус
- Параллельный корпус

2. Компьютерные методы исследования текстов

- тематическое моделирование
- векторные модели семантики

3. Электронное издание текстов

Каталог цифровых ресурсов для башкирского языка

[https://github.com/nevmenandr/bashkir-
language-resources](https://github.com/nevmenandr/bashkir-language-resources)

Каталог языковых ресурсов башкирского

Здесь собирается каталог ссылок на полезные языковые ресурсы башкирского языка

Корпуса башкирского языка / Башкорт теле корпустары

- [Национальный корпус башкирского языка](#), ([инструкция, как искать в корпусе](#))
- [Русско-башкирский параллельный корпус](#)
- [Башкирский поэтический корпус](#), ([как пользоваться корпусом](#))
- [Устный корпус башкирского языка дер. Рахметово и с. Баимово](#)
- [Текстовая коллекция](#)
- [Машинный фонд башкирского языка](#)

Морфология

- [Морфологический анализатор](#)
- [Исемдәрҙең килеш менән үзгәреше / Склонение имен существительных по падежам. \(Онлайн программа\)](#)
- [Морфологический анализатор Урало-Поволжского тюрки](#)

Словари / һүзлектәр

- [Glosbe](#)

Корпуса башкирского языка

Корпус — это

электронное собрание текстов на данном языке с возможностью поиска слов и грамматических конструкций, см. <http://nevmenandr.net/personalia/ВРК.pdf>

Далее будут примеры из НКРЯ (Национальный корпус русского языка, 2005)

- ruscorpora.ru

и Национального корпуса башкирского языка (2019)

- bashcorpus.ru

Основной корпус

Поиск точных форм ?

Слово или фраза

Лексико-грамматический поиск ?

Слово ?

Доп. признаки ? [выбрать](#)

Грамм. признаки ? [выбрать](#)

Расстояние: от до ?

Слово ?

Доп. признаки ? [выбрать](#)

Грамм. признаки ? [выбрать](#)



НАЦИОНАЛЬНЫЙ КОРПУС БАШКИРСКОГО ЯЗЫКА

Стартовая страница башкирского корпуса.

[ПОДРОБНЕЕ](#)

[К КОРПУСУ](#)

bashcorpus.ru

Башкирский корпус

Запрос

● Слово №1

Слово:

Лемма:

Грамматика: 

Глоссы: 



Полнотекстовый поиск:

точное соответствие

Поиск предложений

Поиск слов / лексем



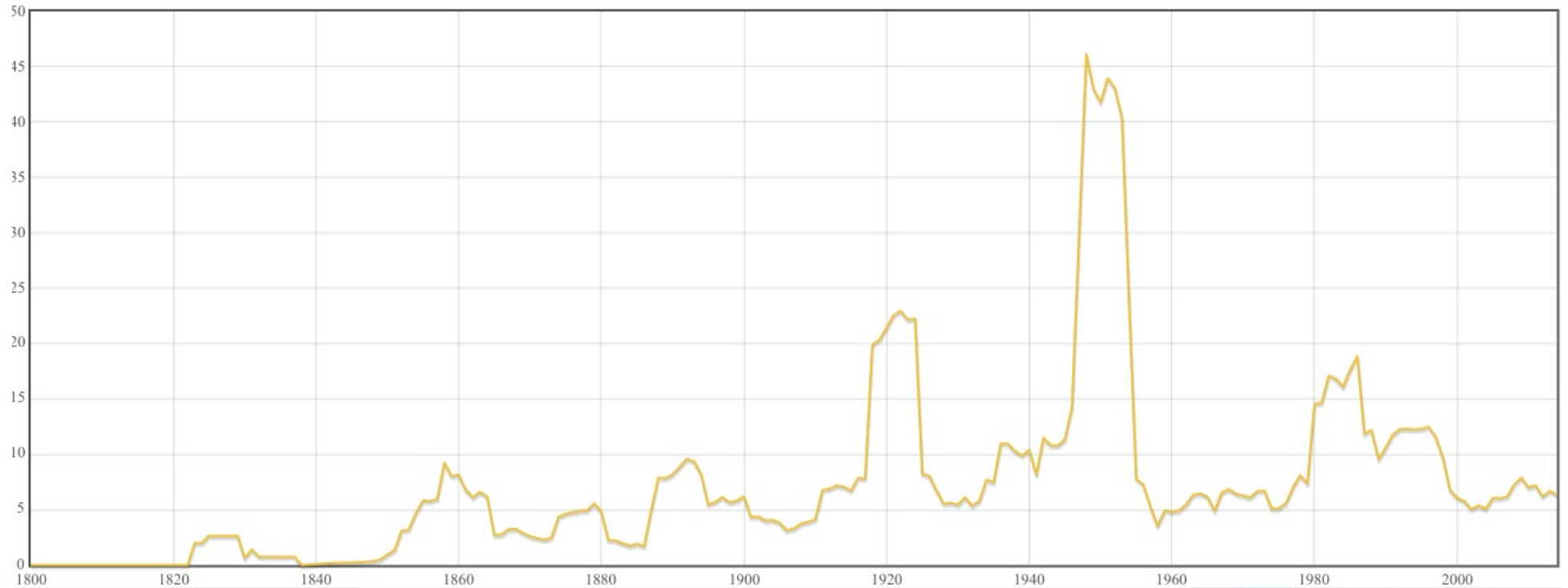
Выбор подкорпуса

(Тут будут результаты.)

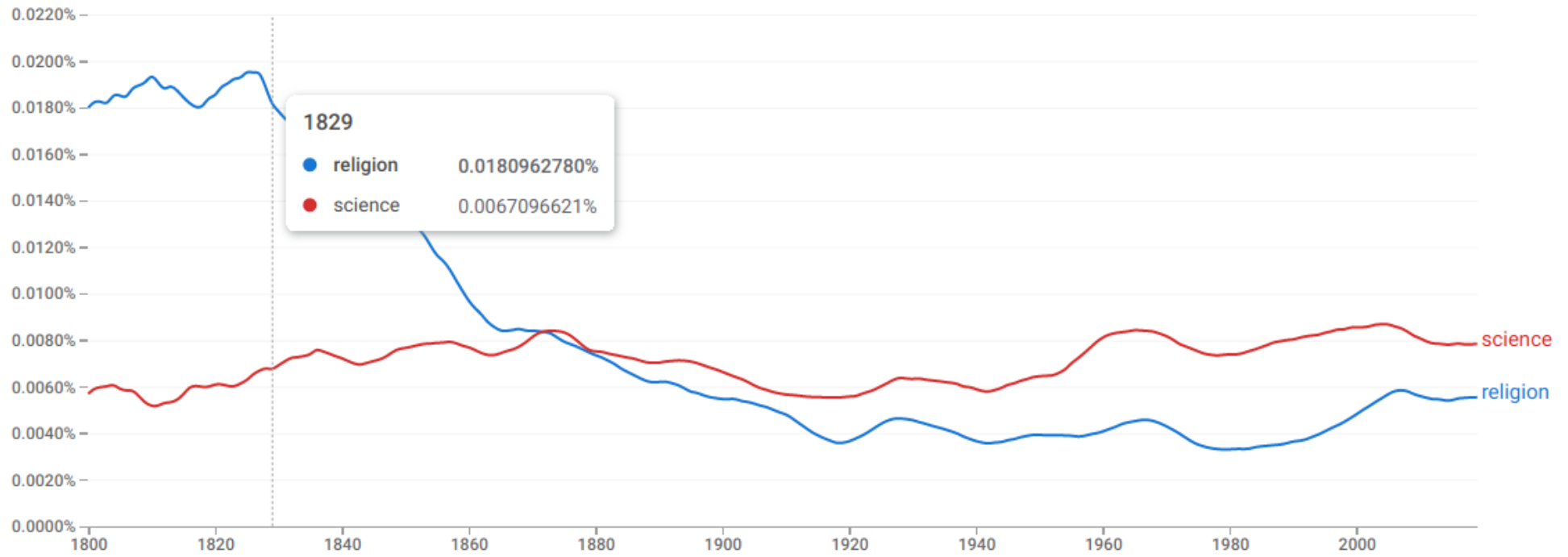
Слово декабрист в русском языке

декабрист

Годы от до со сглаживанием



Культуромика



Языковые варианты

Объем всего корпуса: 115 645 документов, 23 803 881 предложение, 2

явление

Найдено 5 587 документов, 15 565 вхождений.

[Распределение по годам](#) [Статистика](#)

Поискать в других корпусах: [акцентологическом](#), [газетном](#), [диалектно](#)

Страницы: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [следующая страница](#)

1. [коллективный. Форум: Блэйд \(трилогия\) Blade \(2008-2010\)](#) [

[Voland, nick] Сильные сцены: смерть продажного полице
Старейшин... [коллективный. Форум: Блэйд (трилогия) Blade (2008-

Объем всего корпуса: 115 645 документов, 23 803 881 предложение,

явление

Найдено 31 документ, 43 вхождения.

[Распределение по годам](#) [Статистика](#)

Поискать в других корпусах: [акцентологическом](#), [газетном](#), [диалектн](#)

Страницы: [1](#) [2](#) [3](#) [4](#) [следующая страница](#)

1. [А. И. Куприн. Гранатовый браслет \(1911\)](#) [омонимия снята]

Твоя прекрасная нога — **Явление** страсти неземной! Во


Холодильник или һыуыткыс?


Запрос

● Слово №1


Слово:

Лемма:

Грамматика: 

Глоссы: 

Полнотекстовый поиск: точ

Поиск предложений Поиск слов / лексем 

Результат поиска: найдено 2 словоформы, 2 предложений примерн

[Яузарза сыныткан сатирик](#) [Рәдиф Тимершин](#) 2015


Телевизор, **холодильникты** һин, Ялан аяк һыпыртамын.


Запрос

● Слово №1




Слово:

Лемма:

Грамматика: 

Глоссы: 

.....

Полнотекстовый поиск:

точное соответствие

Поиск предложений

Поиск слов / лексем



Выбор подкорпуса



Результат поиска: найдено 431 словоформа, 425 предложений примерно в 378 документах.

Кан баһымын түбәнәйтеү кәрәкһә...

2013



Електәрзе йыуып киптерергә, азақ бер тәүлеккә **һыуыткыска** куйып туңдырырға йәки 1-2 минут кайнатып миләш

Балта Венер Исхаков

2017



Телевизорзан кызыклы кино бара ине. Кара, үззәре көндөзөн дә капкаһын бикләп ултыра. – **һыуыткыста** калғансы...


Вентилятор или өрзөргөс?


Запрос

● Слово №1

Слово:

Лемма:

Грамматика: 

Глоссы: 

Полнотекстовый поиск: точное соответствие

Поиск предложений

Поиск слов / лексем



Выбор подкорпуса

Результат поиска: найдено 2 словоформы, 2 предложений примерно в 2 документах.

Юлдарза үткөн гүмер Рим Исхаков 2016

- Юлда туктап ултырган машина водителе бер сак тэгөрмөскө тын **өрзөргөс** hora булып сыкты.

Йылмайып келөп кенә 2013


- Алдыңғы карашлы ғаилә башлығы ял урынына тура бармай. Кайнап торған сөй карай йүнөлтегез. һауа етмөгән! — тип кушыла өсөнсөһө.


Запрос

● Слово №1




Слово:

Лемма:

Грамматика: 

Глоссы: 

.....



Полнотекстовый поиск:

точное соответствие

Поиск предложений

Поиск слов / лексем



Выбор подкорпуса



Результат поиска: найдено 16 словоформ, 16 предложений примерно в 13 документах.

Киләсәккә илтә канаттар [Е. Моторина](#) 2015

Авиация сәнәгәтенең күп предприятиеларынан айырмалы рәүештә, Өфө моторзар эшләү берекмәһе үз яңы төр изделиелар за сығара алды. Өлеге мәлдә бында айырыу корпусын, эсе кыуыш **вентилятор** ке компрессорының роторын һәм кайһы бер башка деталдәрзе етештерәү үзләштерелде. Өфө заводы бы тәүгеләрзән булып эш башланы.

Более сложные запросы

За

⚙

Лемма:	<input type="text" value="шат"/>	Лемма:	<input type="text"/>
Грамматика:	<input type="text"/>	Грамматика:	<input type="text" value="post"/>
Глоссы:	<input type="text"/>	Глоссы:	<input type="text"/>
Перевод (ru):	<input type="text"/>	Перевод (ru):	<input type="text"/>
Разборы:	<input type="text" value="любое количество"/>	Разборы:	<input type="text" value="любое количество"/>
Позиция в предложении:	<input type="text" value="1"/>	Позиция в предложении:	<input type="text" value="2"/>

Толнотекстовый поиск: точное соответствие

Поиск предложений

Поиск слов / лексем



Выбор подкорпуса

Результат поиска: найдено 760 словоформ, 19 предложений примерно в 19 документах

Котлайбыз! 2012

Шатлык менән озон ғүмер итеп Тыуаһы таңдар һиңә бары тик шатлык-кыуаныс

Батша васыяты Тайсынова 2010

Шатлык менән күңелһезлектең хақын белгән кеше генә үлемдән дә, үлемһезле

“Айгиз, һин иҫәнме ни?!” Айгиз Баймөхәмәтов 2018

Шатлығым менән уртақлашам тип Азат ағайыма шылтыраттым:

Как развивать корпус?

- Нужно **добавлять** новые тексты
 - разного времени
 - разных жанровая
- Нужно исправлять ошибки

Грузия Рәсәйзең терәге булырмы? 2012

Михаил Саакашвили рәттән ике **мөззәт** президент вазифа
һайлауында еңеп сығып, Хөкүмәт башлы

Үз-үзеңә баһа биреп буламы?

«Ике **мөззәт** рәттән» мәсьәләһендә Кон
белдерзе.

мөззәт
мөззә? v
мөззә?-т
gr: caus
trans_ru: ?

Параллельный корпус

<https://ruscorporora.ru/new/search-para-ba.html>

2. Мостай Кэрим. Озон-озақ бала саК (1976) | Мустай Карим. Долгое-долгое детство (И. Каримов, 1977)

ru	Как тени двух лун , будет во мне двойная грусть от их двойной любви. [Мостай Кэрим. Озон-озақ бала саК ←...→
ba	Ике мөхәббәттең шәуләе булып гел генә ике Һағыш эйәрәсәк миңә. [Мустай Карим. Долгое-долгое детство (
ba	Без генә түгел, ошо имән тирәләй бөтә ер, бөтә күк әйләнә. Кояш, ай, йондоҙҙар әйләнә. Ә ағас Җым Долгое-долгое детство (И. Каримов, 1977)] [омонимия не снята] ←...→
ru	И не только мы кружится земля вокруг дуба, небо. Солнце кружится, луна , проглядывающие звезды. Каримов, 1977)] [омонимия не снята] ←...→
ba	Ай күрзе, Кояш алды... Минең күңелемә иләҫ-миләҫ уйҙар Җағылып үткәнде дуҫым һиҙмәне. [Мос 1977)] [омонимия не снята] ←...→
ru	Луна видела, солнце высушило... Друг не заметил, что я о своем задумался. [Мустай Карим. Долгое-долго

Как развивать корпус:

Можно **выравнивать** новые тексты и добавлять их в проект

Как использовать корпус:

Как **словарь**, только лучше, потому что

- есть перед глазами контекст
- поиск ведется по всем формам

Компьютерные методы исследования текстов

Борис Валерьевич Орехов – доцент Школы лингвистики Национального исследовательского университета «Высшая школа экономики».

Представлены результаты всестороннего количественного исследования башкирской системы версификации в XX веке. С использованием статистических инструментов подвергаются анализу все уровни организации стихотворного текста от фонетики до лексики и грамматики, с особым вниманием к метру и ритму. Количественные данные получены на корпусе текстов 103 башкирских поэтов общим объемом в 1,77 млн. словоупотреблений. Анализ предварен подробным обзором науки о тюркском стихе, начиная с 1950-х годов. Основную роль в башкирском стихосложении XX века играют силлабические формы фольклорного происхождения узун-кюй и кыска-кюй, первая из которых специфична для поволжско-кыпчакского слогосчитающего стиха. Приводится подробное сопоставление башкирского стиха с киргизским. Книга завершается примерами поэтических текстов на башкирском языке, сгенерированных с использованием искусственных нейронных сетей.



Башкирский стих XX века

Б. В. ОРЕХОВ

Б. В. ОРЕХОВ

Башкирский стих XX века

КОРПУСНОЕ
ИССЛЕДОВАНИЕ

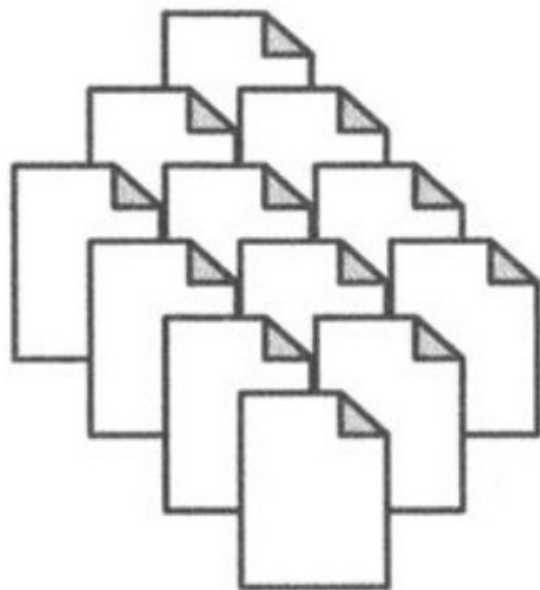


Башкирский стих XX века (2019)

Тематическое моделирование

- Возможность с применением компьютерных методов узнать тематику текста, **не читая** его.
- Имеет смысл, когда текста **слишком много** для человека.
- Анализируя **совместную встречаемость** слов в текстах, компьютер показывает, какие темы эти слова образуют.

Документы из корпуса
включают комплекс тем



Тема — это
распределение слов

айоли обжарить блины
кобе крем чеснок

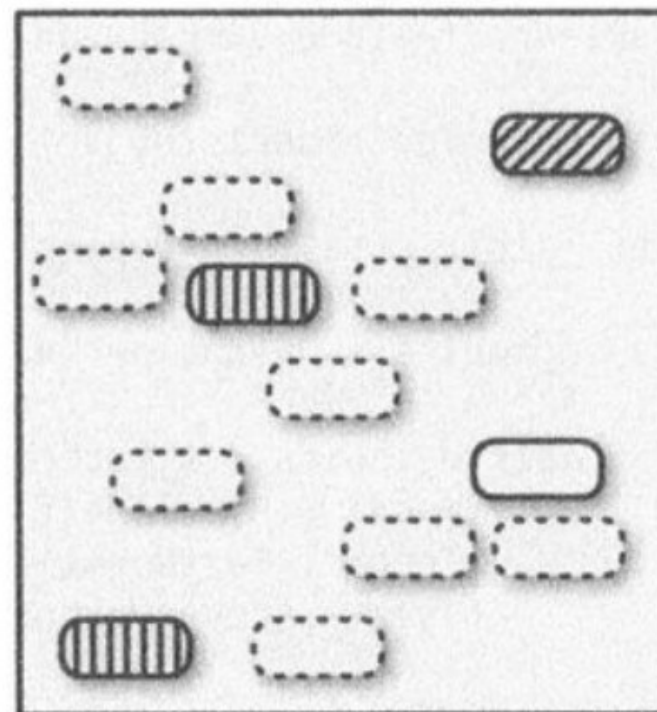
выборы печать бежать
президент иракский

кобе НБА рекорд
план финал сухожилие

примечание запись поп
виолончель гармония трио

честный консоль игра
единство тактический

Один документ охватывает
несколько тем



В башкирской поэзии XX века

Для 1960-х годов получился следующий набор тем:

Тема 1: *юл* 'дорога', *кил* 'приходить', *кеше* 'человек', *йөрәк* 'сердце', *күз* 'глаз', *күр* 'смотреть', *тор* 'стоять', *көн* 'день', *кит* 'уходить', *юк* 'отсутствие'.

Тема 2: *ти* 'говорить', *әйт* 'сказать', *һүз* 'слово', *егет* 'парень', *карт* 'старый', *икән* 'быть', *эш* 'труд', *эс* 'пить', *бел* 'знать', *дуҫ* 'друг'.

Тема 3: *кар* 'снег', *кыш* 'зима', *яз* 'весна', *яу* 'литься', *буран* 'буран', *һалкын* 'холодный', *сәскә* 'цветок; волосы', *ак* 'течь; белый', *һағын* 'осторожный, бдительный, бережливый', *март*.

Тема 4: *йыр* 'песня', *моң* 'печаль, напев', *һандугас* 'соловей', *һайра* 'петь', *йырла* 'петь', *кош* 'птица', *күңел* 'душа, чувства', *шишмә* 'родник', *яз* 'писать; весна', *тыу* 'знамя'.

Тема 5: *тау* 'гора', *диңгез* 'море', *тулкын* 'волна', *кая* 'скала', *таш* 'камень', *йылға* 'река', *бейек* 'высокий', *Урал*, *ағ* 'течь', *яр* 'голос'.

Тексты 1980-х годов дали такие темы:

Тема 1: *кил* 'приходить', *күңел* 'душа, чувства', *йөрәк* 'сердце', *ти* 'говорить', *донъя* 'мир', *юл* 'дорога', *күр* 'смотреть', *бар* 'идти', *юк* 'отсутствие', *кит* 'уходить'.

Тема 2: *йәш* 'молодой', *һайра* 'петь', *уз* 'проходить, миновать', *бел* 'знать', *гүмер* 'жизнь', *сәскә* 'цветок; волосы', *йәше* 'спрятать', *һиз* 'чувствовать', *гашик* 'влюбленный', *хушлаш* 'проститься'.

Тема 3: *тыу* 'знамя', *як* 'местность', *таң* 'заря', *тупрак* 'земля, почва, грунт', *гөл* 'цветок', *төйәк* 'родина', *сәскә* 'цветок', *моң* 'печаль, напев', *үс* 'расти', *йондоз* 'звезда'.

Тема 4: *япрак* 'лист', *йәшел* 'зеленый', *урман* 'лес', *көз* 'осень', *кар* 'снег', *ак* 'белый', *кайын* 'береза', *төс* 'цвет', *һары* 'желтый', *кой* 'осыпаться'.

Тема 5: *кыз* 'девушка', *ай* 'луна', *тор* 'стоять', *гашик* 'влюбленный', *ак* 'белый', *ысынла* 'истинный', *аксарлак* 'чайка', *бала* 'дитя', *кара* 'черный', *гөл* 'цветок'.

Векторные модели семантики

- Используя определенную технологию, можно **вычислять семантику** слова, основываясь только на текстах.
- Не нужно составлять **словарь** (почти).
- Семантические связи слов можно представить в виде **карты**.

Для русского языка

<https://rusvectors.org/ru/>

Семантические ассоциаты для **разум** (ALL)

Частотность слова

Высокая Средняя Низкая

НКРЯ и Wikipedia

1. рассудок NOUN 0.78
2. ум NOUN 0.65
3. здоровый ADJ 0.64
4. разум PROPN 0.62
5. разумение NOUN 0.62
6. сознание NOUN 0.61
7. истина NOUN 0.61
8. познание NOUN 0.60
9. мудрость NOUN 0.59
10. мышление NOUN 0.59



Для башкирского языка

Модель башкирского поэтического корпуса (по
мы указано значение косинусной близости между *йорт*

нигез 'фунамент' 0,762615

өй 'дом' 0,756706

ауыл 'деревня' 0,690155

тупһа 'крыльцо' 0,625492

бина 'здание' 0,616340

яг 'сторона' 0,601330

өфө 'Уфа' 0,596297

гостиница 'гостиница' 0,594483

өстәл 'стол' 0,594145

мөйөш 'угол' 0,586830

Модель корпуса газеты «Йэшлек»:

бина 'здание' 0,664548

һарай 'дворец' 0,579985

бүлмә 'комната' 0,543696

өй 'дом' 0,536694

мәсет 'мечеть' 0,515221

интернат 'интернат' 0,505365

бакса 'огород' 0,504052

дауахана 'больница' 0,500521

ятак 'ночлег' 0,4967596

арай 'пойма' 0,4957030

Электронное издание

Фольклорный архив (2011)

<http://nevmenandr.net/pages/bashfolk.php>
(сайт на базе БашГУ уничтожен)

Фольклорный архив Башкирского государственного университета

ISBN 978-5-87604-352-8

- [Оглавление издания.](#)
- [Описание издания.](#)
- [Указатель словоформ.](#)
- [Карта районов записей.](#)

Описание издания

История собирательства, публикаций и изучения фольклора в Республике Башкортостан составляет без малого две сотни лет. Огромный вклад в деятельность по сбору фольклорных произведений внесли такие ученые и деятели культуры как С. Рыбаков, М. Уметбаев, А. Бессонов, М. Лось, Д. Зеленин, Г. Тукай, М. Гафури, Ф. Туйкин, З. Уммати, Ш. Бабич, А.-З. Валиди, А. Инан (Ф. Сулейманов), М. Бурангулов, Х. Габитов, С. Мирас, Г. Вильданов, Д. Юлтый, Г. Амантай, Г. Саям, А. Карнай, Г. Амири, А. Харисов, А. Усманов, Л. Лебединский, М. Салтыков, К. Мэргэн, М. Сагитов, С. Галин, Л. Бараг, М. Мингажетдинов, Ф. Надршина, Н. Шункаров, Н. Зарипов, А. Сулейманов, Б. Ахметшин, М. Мамбетов, Р. Султангареева, Р. Шакуров, Б. Баимов, Л. Атанова, Р. Сулейманов, Х. Ихтисамов, Ф. Камаев, Л. Сальманова, Г. Хусаинов, И. Галяутдинов, Г. Хусаинова, А. Кубагушев, Ю. Гайнетдинов, Н. Ахметжанова, Г. Галина, Ф. Фаизова, Ф. Ахатова.

Интерактивное оглавление

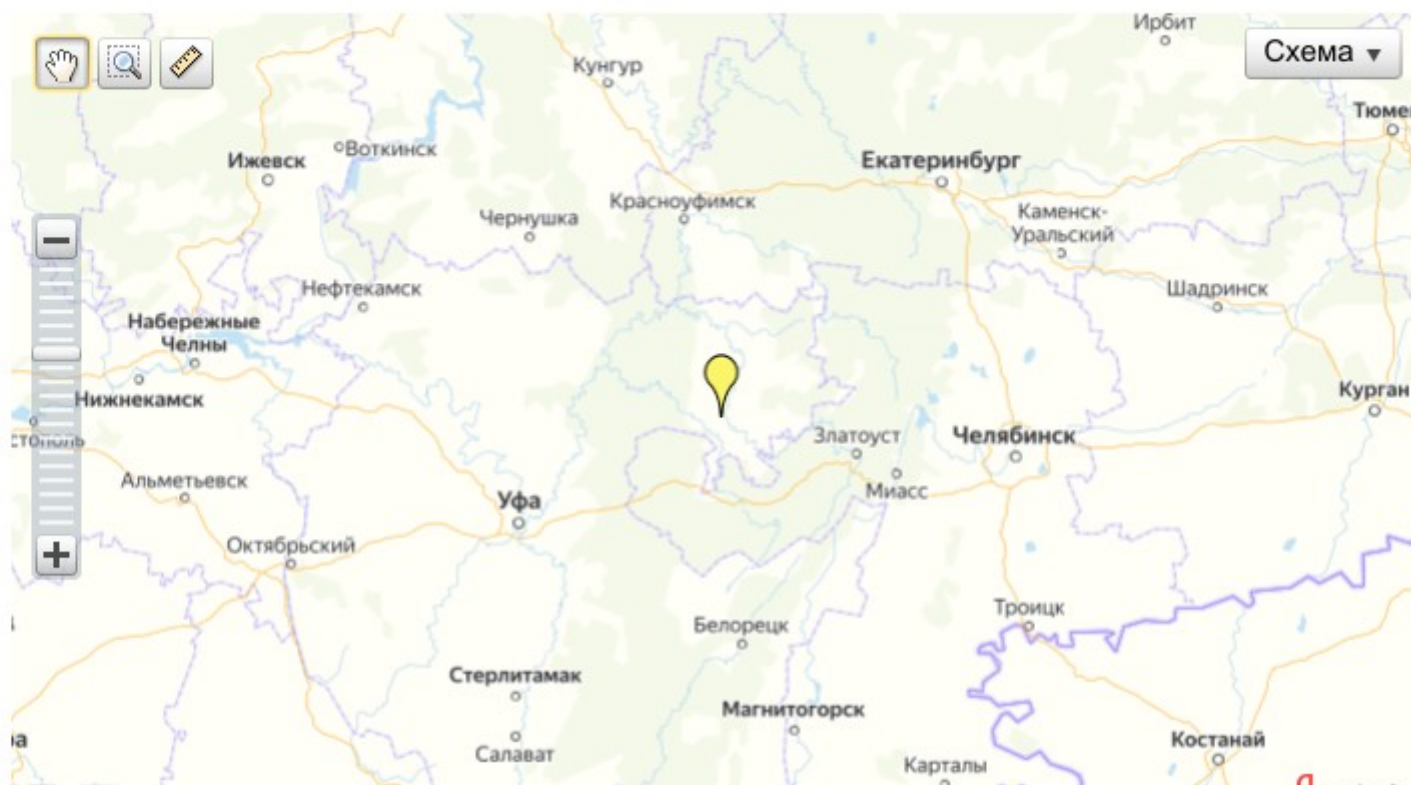
Фольклорный архив Башкирского государственного университета

Оглавление издания						Фильтр
ID	Название	Жанр	Район	Год записи	Ссылка	
1	Таш һын	сказка в	Салаватский район	1963	д. 1, л. 3—10	
2	Ғәлиә	сказка в	Салаватский район	1963	д. 1, л. 11—16	
3	Энем һары ат	сказка в	Салаватский район	1963	д. 1, л. 17—20	
4	Кеҫә-моҫа	сказка б	Салаватский район	1963	д. 1, л. 21—23	
5	Тәкәббер ҡатын	сказка б	Салаватский район	1963	д. 1, л. 27	
6	Өкиәт	сказка в	Салаватский район	1963	д. 1, л. 31—33	
7	Дәйә һырты	сказка б	Салаватский район	1963	д. 1, л. 24—26	

Карта с отметкой о месте записи текстами

— Убыр үлһә, кара кан сығыр, без үлһәк, кызыл кан сығыр. Шулай тинеләр зә, улыр менән һыу эсендә һугыша башланылар, ти, былар тешлөгәйне, кызыл кан ургылып килеп сыкты, ти. Егет хуркып, кыскырып ебәрә яззы, ти.

Күп тә үтмәй һыу караған булып кайнай башланы, ти. Озақ вахыт үтмәне, һыузан эттәр зә сыктылар, өсәүләп шатланышып кайтып кит



Указатель словоформ

Указатель словоформ

[← Предыдущая страница](#) • [Следующая страница](#) → | [Частотный словарь](#)

[\[1\]](#) [\[2\]](#) [\[3\]](#) [\[4\]](#) [\[5\]](#) [\[6\]](#) [\[7\]](#) [\[8\]](#) [\[9\]](#) [\[10\]](#) [\[11\]](#) [\[12\]](#) [\[13\]](#) [\[14\]](#) [\[15\]](#) [\[16\]](#) [\[17\]](#) [\[18\]](#) [\[19\]](#) [\[20\]](#) [\[21\]](#) [\[22\]](#) [\[23\]](#) [\[24\]](#) [\[25\]](#) [\[26\]](#) [\[27\]](#) [\[28\]](#) [\[29\]](#) [\[30\]](#) [\[31\]](#) [\[32\]](#) [\[33\]](#) [\[34\]](#) [\[35\]](#) [\[36\]](#) [\[37\]](#) [\[38\]](#) [\[39\]](#) [\[40\]](#) [\[41\]](#) [\[42\]](#) [\[43\]](#) [\[44\]](#) [\[45\]](#) [\[46\]](#) [\[47\]](#) [\[48\]](#) [\[49\]](#) [\[50\]](#) [\[51\]](#) [\[52\]](#) [\[53\]](#) [\[54\]](#) [\[55\]](#) [\[56\]](#) [\[57\]](#) [\[58\]](#) [\[59\]](#) [\[60\]](#) [\[61\]](#) [\[62\]](#) [\[63\]](#)

№	Словоформа	Частотность
5101	итәкәйзәрен	1
5102	итәлмам	1
5103	итәләр	18
5104	итәм	19
5105	итәмен	1
5106	итәр	4
5107	итәргә	2

Контексты (конкорданс)

Указатель словоформ. Контексты слова «итәләр»

Таш һын (сказка (волшебная)), 1963 г.

...а. Биш минутка һөйләргә рөхсәт **итәләр**. Батша малайына ниндәй яхшылык...

Бай (Сказка (бытовая)), 1964 г.

...н була. Дәһлә был ярлыны хөкөм **итәләр**. һорансы судта: ...

<Өкиәт> (Сказка (бытовая)), 1969 г.

...ка һалып, һыуға ташларға карар **итәләр**. Бермимылдыкты токка һалалар з...

Кара каз (Сказка (волшебная)), 1945 г.

...бикәгә ярай. Батша улына хәбәр **итәләр**. Гөлбикә агас кыуышына барып, ...

Өрмәнде тун (Сказка (волшебная)), 1957 г.

...ик татыу ғына йөшөүзәрен дауам **итәләр**, ти. ...

Приятные мелочи

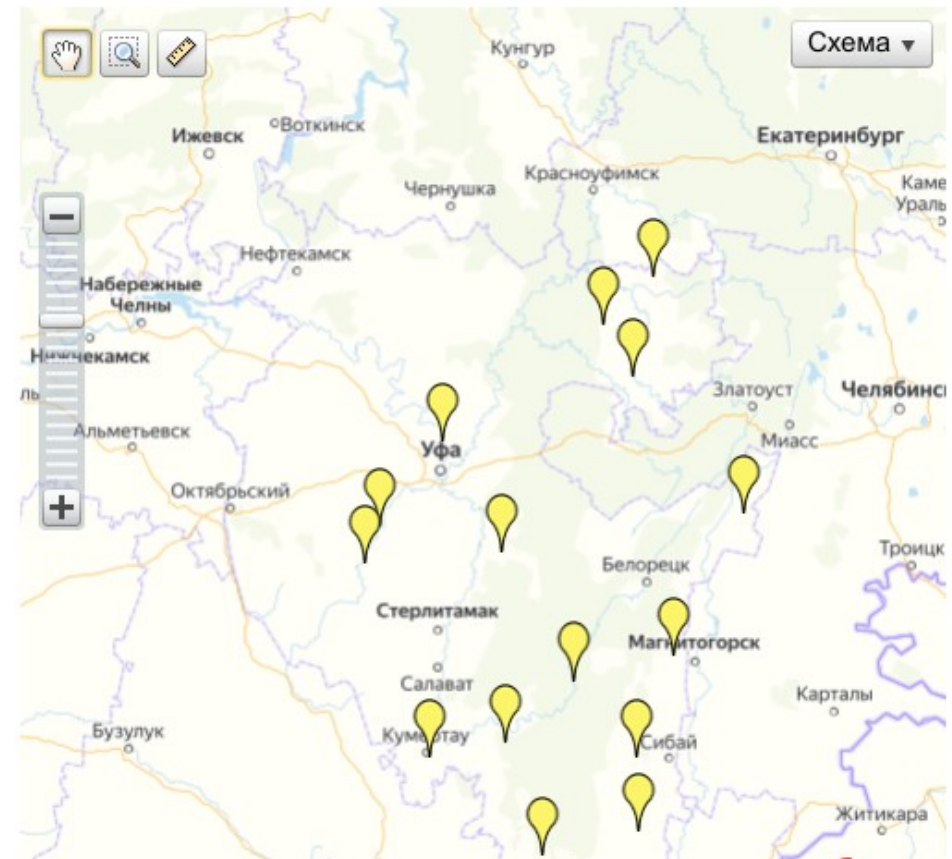
— Әһә, миңең катыңды үбергә итә, — тип, иртәгәһен б

Егетте бейек тау башына атырға алып менәләр. Халык минутка һөйләргә рөхсәт **итәләр**. Батша малайына ни бил, муйын, азақ килеп бөтөнләйгә ташка әйләнгән дә илагандар икән, ти.

Ссылка на эту публикацию:

Таш һын [Электронный документ] // Фольклорный архив
Б. В. Орехова, А. А. Галлямова. [2011—2021]. Дата отбора
id=1&w=%D0%B8%D1%82%D3%99%D0%BB%D3%99

Картография опубликованных в ЭНИ фольклорных записей:



Ссылаются ли специалисты?

Муратова Р.Т. Словарь-конкорданс башкирских народных песен о Родине. - Уфа, 2016. С. 13.

В последние годы имеется возможность составления конкордансов с помощью национальных корпусов. В этом плане можно отметить Корпус башкирской прозы, Корпус башкирской публицистики [<http://212.193.132.98/bashcorp/bashcorp>, <http://212.193.132.98/bashcorp/korpuspub>]. Разрабатывается корпус башкирского фольклора.

В общем, составление конкордансов не является новым в башкирской лексикографии. А вот конкорданс башкирских народных песен о Родине является первым опытом создания конкордансов по тематическому песенному фольклору.

Что нужно делать?

Создавать другие современные электронные издания значимых для башкирской культуры текстов

Moore Archive EDITIONS AND RESOURCES EDUCATION AND SCHOLARSHIP ABOUT

Home / testing

Thumbs Image-Text Show All Links

07-03-11 0021-verso

Key: Unclear Text Deleted Illegible Deletion: Illi

ever inherited

capital or earned capital or used

shld be denied the right

capital should be is disqualified to

enses

hold office, but you might as well

vote this year for the peanut king

(whose brain has been removed &

lent to Columbia as referee in the polit

economy department—) who cannot

debate a matter for or against—who

just rules. Seriously—the country would

anyone

Eidólons.
(printed in [Two Rivulets](#) 1876)

I met a Seer,
Passing the hues and objects of the
world,
The ? sweets ^{2fruits} ^{2joys} ^{2fields} ^{2fruits} of art, material joys & learning, pleasure,
& sense
To glean Eidólons
Put in thy chant, said he.
No more the visible human ^{fleeting, fractional} face
or limb,
Nor hour, nor day—no ^{segments}, parts put in,
Put in Eidólons only.

Lo! I or you,
Or ^{Woman}, Man ^[about three letters deleted, illegible] or State, known or unknown
While ^{We} seeming solid ^{beauty} wealth, er
strength ^{ef beauty} we build, we ² better build;
We ^{But really} build Eidólons.

Что требуется для развития компьютерных методов изучения башкирского языка?

- **Оцифровка!** Все инструменты искусственного интеллекта строятся на данных (текстах)
- Участие экспертов-носителей языка в **исправлении ошибок.**
- Всё это можно встроить в образовательный процесс.

