

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

файдаланучылар өчен кулланма

v3.4

Мансур Сайхунов

Тәүзих Ибраһимов

Jorma Luutonen

Содержание

- [Татар теленең язма корпусы турында гомуми мәгълүмат](#)
- [Тел корпусы – ул нигә кирәк?](#)
- [Татар теленең язма корпусы ничек баетылачак?](#)
- [Системага таләпләр](#)
- [Татар теленең язма корпусына тупланган мәгълүмат белән танышу](#)
- [“Эзләү” битенә кире кайту](#)
- [Татар хәрефләрен ничек язарга белмәсәгез!](#)
- [Эзләү нәтижәләренә туры сылканы алу](#)
- [Корпуста аерым сүзформа һәм аның контекстына бәйлә статистик мәгълүматны табу](#)
- [Корпуста аерым лексеманы табу](#)
- [Корпуста аерым префикска башланган сүзләрне табу](#)
- [Корпуста билгеле бер хәрефләр жыелмасына төмамланган сүзләрне табу](#)
- [Корпуста кирәкле сүзтезмәне табу](#)
- [Катлаулы морфологик эзләү системасы:](#)
 - [Өлеге системаның асылы турында](#)
 - [Корпуста сүзформаларны үз эченә алган комбинацияләр нигезендә эзләү](#)
 - [Корпуста леммаларны үз эченә алган комбинацияләр нигезендә эзләү](#)
 - [Корпуста морфологик \(грамматик\) тегларны үз эченә алган комбинацияләр нигезендә эзләү](#)
 - [Корпуста сүзнең башы, уртасы һәм/яки азагы күрсәткече кулланылган комбинацияләр нигезендә эзләү](#)
 - [Корпуста эзләнәсе нәр сүзгә берничә параметр күрсәтү](#)
 - [Эзләү параметрларын график режимда кертү](#)
- [Мисаллар буларак китерелгән жөмлөләрне тыңлап карау](#)
- [Өстәмә статистик мәгълүмат!](#)
- [Басылган мәкаләләр](#)
- [Корпус тасвирламасы](#)
- [Корпустан файдалану](#)
- [Проект турында](#)
- [Рәхмәт сүзе](#)
- [Элемтәгә керү өчен мәгълүмат](#)
- [Башка телләрдә](#)

Татар теленең язма корпусы турында гомуми мәгълүмат

Әлеге корпус хәзерге язма татар телен мөмкин кадәр тулырак чагылдыруга юнәлгән электрон формадагы чыганак. Бүгенге көндә корпусның күләме 116 млн. сүздән артык, төрле сүзформаларның саны 1 млн. 500 меңгә кадәр житә. Корпусны төзүдә Интернеттагы төрле ачык ресурслардан алынган материаллар кулланылды. Текстларның яртысынан артыгы (Корпусның гомуми күләменнән $\approx 60\%$) татар теленең публицистик стилиен хасил итә. Язма әдәби татар теленең башка стильләре Корпуста түбәндәге өлештә урын ала:

- матур әдәбият стиле $\approx 35\%$,
- гуманитар профильдәге фәнни әдәбият стиле $\approx 4\%$,
- эш кәгазьләре стиле $\approx 1\%$.

Лингвистик материалның төп чыганаклары буларак төрле WEB-сайтларда урнаштырылган текстлар алынды, мәсәлән, “Татар – информ”, “Ислам – татарлар һәм мөселманнар”, “Татар электрон китапханәсе”, “Дуслык”, “Азатлык”, “Яңа гасыр”, “Татар әдипләре” һ.б.

Татар теле язма корпусының төп максаты – татар теле лексик-семантик системасын статистик лексикология һәм когнитив лингвистика кысаларында тирәнрәк өйрәнүгә юл ачу.

Корпус өч режимда кулланылырга мөмкин:

- статистик (аерым бер сүзгә табу, аның уң, сул, семантик контекстларын билгеләү, сүзформаларның кулланылыш ешлыгын ачыклау),
- лемматизация (морфологик разметка булу сәбәпле, теге яки бу леммага караган сүзформаларны эзләү, аларга хас грамматик категорияләргә билгеләү),
- сүзформаны шаблон буенча эзләү (сүзнең башы, уртасы һәм/яки азагы нигезендә туры килә торган очракларны ачыклау).

Моннан тыш, корпусның эзләү системасы аша табылган жөмлөләрне яңгыратып карау (тыңлау) мөмкинлеге дә каралган. Кирәк булса, корпус-менеджер уң һәм сул контекстларны киңәйтеп күрсәтергә мөмкин, ягъни бу сүзнең яки сүзтөзмәнең мәгънәсен ачыклаудагы төгәлләкне арттыра.

Язма корпусның эзләү системасы түбәндәге операцияләргә башкарырга мөмкинлек бирә:

- кирәкле сүзләргә эзләү, аларның кулланылу ешлыгын билгеләү, бу сүзнең яки форманың телдә кулланылышын дәлилләүче мисаллар табу;
- кайсы сүзләр эзләнәсе сүзнең алдында, аңардан соң һәм, гомумән, бер жөмлө кысаларында килергә мөмкинлекне ачыклау (әлеге сүзнең сул, уң, семантик контекстлары), шулай ук, әлеге сүзләрнең эзләнәсе сүз белән ни дәрәжәдә еш кулланылуын ачыклау һ.б.

Тел корпусы - ул нигә кирәк?

Тел – шушы чор мәдәниятенә көзгесе. Халыкның нинди шартларда, нинди уйлар белән яшәве аның телендә чагыла. Тәкъдим ителгән Татар теленә Язма корпусы бүгенге көн һәм XX гасырның икенче яртысына хас чәчмә матур әдәбият стилин, фән, дин тәгълиматы, эш кәгазьләре телләрен адекват гәүдәләндерә. Язма корпусның 60% чамасы WEB-сайтлар нигезендә төзелгән булуы алдагы буыннарда шушы чорда яшәүчеләрнең уй-фикере, тормыш шартлары белән танышу мөмкинлекләре тудыра. Аның зур сандагы текст һәм жөмлөләргә нигезләнгән булуы, шуңа өстәп төрле эзләү программалары белән җиһазландырылуы бүгенге телчеләрне татар телен концептология, культурология, когнитология кысаларында өйрәнү мөмкинлекләре белән тәэмин итә.

Рус халкы мәкалендә: “Аппетит приходит во время еды» - дип әйтелә. Хәерле сәгатьтә!

Татар теленең язма корпусы ничек баетылачак?

Лингвистика бүгенге көндә күп тармакларга бүленеп тел өйрәнүне антропоцентризм юнәлешендә дәвам итә. XX гасыр тел белемендә тирән эз калдырган телчеләрнең берсе - Леонард Блумфильд үзенең “Тел” дип исемләнгән монографиясен: “Хәзергә хыял булып кына күренсә дә, ышанырга була ки, телне өйрәнү ерак булмаган киләчәктә безгә кешеләрнең кылган эшләрен аңлауда һәм алар белән идарә итүдә ярдәм күрсәтәчәк” - дигән сүzlәр белән тәмамлый.

Шушы аспектта көн тәртибенә телнең үзе түгел, ә бәлки тел ярдәмендә мәгълүмат тапшыру, алмашу процессын өйрәнү, аның моделен төзү проблемасы куелган. Әлеге планда телнең тезаурус сүзлекләрен, тел үзгәрешен чагылдырган, әдипләрнең индивидуальлеген чагылдырган зур булмаган корпуслар (подкорпуслар) төзү, шулар белән беррәттән корпуслар лингвистикасын тематик яктан баету - алда торган проблемаларның бер өлеше.

Системага таләпләр

Корпус белән эшләр өчен техник киңәшләрәбез:

1. Алда күрсәтелгән браузерларның яңа версияләрен куллану:

- Mozilla Firefox: <https://www.mozilla.org/en-US/firefox/all/#ru>
- Google Chrome: <https://www.google.ru/intl/ru/chrome/browser/desktop/index.html>

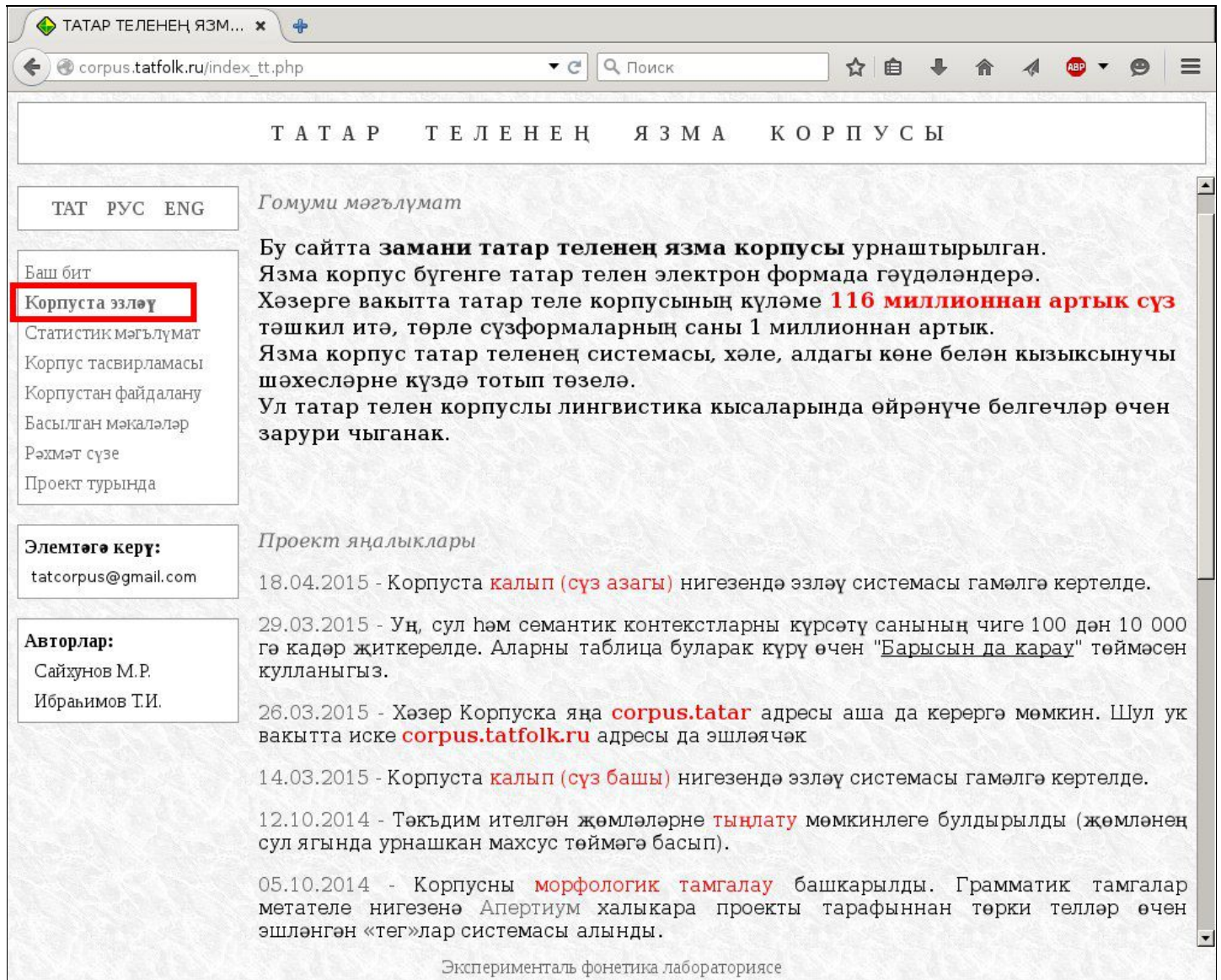
2. Браузерда JavaScript системасын кабызу.

3. Жәмләләрне яңгырату эшләмәсә:

- башка браузер кулланып карагыз,
- Adobe Flash Player'ның яңа версиясен куегыз: <https://get.adobe.com/flashplayer/>

Татар теленең язма корпусына тупланган мәгълүмат белән танышу

Интернетта corpus.tatar (яки corpus.tatfolk.ru) адресында урнашкан сайтка кереп баш битенең сул ягындагы “Корпуста эзләү” ссылкасына чиртегез:



The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main navigation menu includes "ТАГ", "РУС", and "ENG". A sidebar on the left contains a list of links, with "Корпуста эзләү" highlighted in a red box. The main content area is titled "Гомуми мәгълүмат" and contains a paragraph of text. Below this is a section titled "Проект яңалыклары" with a list of dates and news items. At the bottom, there is a footer that reads "Эксперименталь фонетика лабораториясе".

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

Гомуми мәгълүмат

Бу сайтта замани татар теленең язма корпусы урнаштырылган. Язма корпус бүгенге татар телен электрон формада гәүдәләндерә. Хәзерге вакытта татар теле корпусының күләме **116 миллионнан артык сүз** тәшкил итә, төрле сүзформаларның саны 1 миллионнан артык. Язма корпус татар теленең системасы, хәле, алдагы көне белән кызыксынучы шәхесләрне күздә тотып төзелә. Ул татар телен корпуслы лингвистика кысаларында өйрәнүче белгечләр өчен зарури чыганак.

Проект яңалыклары

18.04.2015 - Корпуста **калып (сүз азагы)** нигезендә эзләү системасы гамәлгә кертелде.

29.03.2015 - Уң, сул һәм семантик контекстларны күрсәтү санының чиге 100 дән 10 000 гә кадәр житкерелде. Аларны таблица буларак күрү өчен "Барысын да карау" төймәсен кулланыгыз.

26.03.2015 - Хәзер Корпуска яңа **corpus.tatar** адресы аша да керергә мөмкин. Шулу вакытта иске **corpus.tatfolk.ru** адресы да эшләячәк

14.03.2015 - Корпуста **калып (сүз башы)** нигезендә эзләү системасы гамәлгә кертелде.

12.10.2014 - Тәкъдим ителгән жөмлөләрне **тыңлату** мөмкинлегә булдырылды (жөмлөнән сул ягында урнашкан махсус төймәгә басып).

05.10.2014 - Корпусны **морфологик тамгалау** башкарылды. Грамматик тамгалар метатөле нигезенә Апертиум халыкара проекты тарафыннан төрки телләр өчен эшләнгән «тег»лар системасы алынды.

Эксперименталь фонетика лабораториясе

Чиртү нәтижәсендә файдаланучыга “Татар теленең язма корпусында эзләү” дигән бит ачыла:

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area is titled "Татар теленең язма корпусында эзләү" and contains a search form with a text input field and a "Табу!" button. The search instruction reads: "Эзләү төрен сайлап, кирәкле сүзне языгыз:". Below the search form, there are three radio button options for search criteria: "Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)", "Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)", and "Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *дөргө**)". The left sidebar contains navigation links: "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустаң файдалану", "Басылган макаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтөгә керү: tatcorpus@gmail.com", and "Авторлар: Сайхуннов М.Р., Ибраһимов Т.И.". The footer text is "Эксперименталь фонетика лабораториясе".

“Эзләү” битенә кире кайту

Аерым бер мәгълүмат эзләүне башкарганнан соң, корпуca башка йөкләмә бирү өчен түбәндәге боерыклардан файдаланырга мөмкин. Бу очракта, аланчыкларга керткән барлык мәгълүмат та саклана.

TATAR TELENEŇ ЯЗМА КОРПУСЫ

TAT RUS ENG ["Эзләү" битенә кайту] [Жөмлөләргә күчү]

Баш бит
Корпуca эзләү
Статистик мәгълүмат
Корпуca тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Г.И.

СҮЗ: *китап*
ТАБЫЛУ САНЫ: 31866

УҢ КОНТЕКСТ: нәшриятында (1409; 19717.9), нәшрияты (1026; 14194.5), булып (871; 3345.21), авторы (869; 8785.8), итеп (787; 3081.54), " (628; 620.15), та (555; 2390.1), белән (476; 140.15), дөнья (407; 2037.03), » (392; 435.78), укырга (382; 2465.02), нәшриятының (379; 5576.17), укучылар (305; 1878.74), укып (304; 1943.77), - (303; 43.75), чыгару (262; 1630.57), күргәзмәсе (254; 2204.72), уку (252; 1002.89), нәшер (237; 1964.75), бүлөк (204; 972.04), бастырып (196; 1450.99), турында (184; 205.99), бастыру (184; 1679.97), басылып (173; 1192.09), чыгарган (163; 1126.32), рөвешендә (160; 1110.37), чыгарырга (160; 1137.4), басу (159; 1292.96), укучы (157; 831.22), укый (150; 766.25), нәшриятына (149; 2191.66), алып (144; 133.16), " (143; 12.47), язган (142; 690.19), язып (142; 709.43), сөючеләр (134; 1248.08), — (132; 20.56), язарга (131; 853.71), укымый (127; 1443.4), чыкты (126; 399.29), кибете (126; 1205.3), нәшрияты (125; 2116.87), бар (123; 59.78), чыгара (117; 659.77), чыгарды (113; 743.19), яза (112; 496.18), язу (105; 623.5), укуга (101; 951.18), язарлык (100; 1228.2), сәүдәсе (99; 1153.05), кибетләрендә (99; 1098.99), сатып (98; 307.17), - (97; 4.65), күргәзмәләре (96; 1068.96), кибетендә (94; 1000.95), тәкъдим (92; 169.96), кибетләре (91; 973.71), буларак (81; 153.79), кибетенә (80; 861.09), сүзе (74; 339.54), чыгарып (74; 347.04), фонды (70; 280.96), кына (69; 99.87), сату (65; 288.18), киштәсе (65; 920.46), килеп (63; 36.21), укыган (63; 286.33), өстендә (61; 220.12), авторлары (56; 416.63), киштәләрендә (56; 616.84), кебек (55; 14.85), киштәләре (55; 669.43), тәкъдир (53; 303.83), аша (50; 64.75), теле (50; 89.08), тотып (50; 137.34), кибетләренә (50; 582.79), алырга (49; 86.31), нәшрияте (49; 637.86), чыккан (48; 66.4), укучыларны (48; 277.7), тарату (48; 315.49), алам (46; 201.67), укыйм (46; 313.4), күргәзмә-ярминкәсе (45; 516.95), нәшр (43; 445.79), укуны (42; 332.42), яки (41; 34.96), укучыларга (40; 194.98), нәшриятлары (40; 490.78), укучыларнын (39; 169.41), язды (38; 206.85).

Эксперименталь фонетика лабораториясе

Шулай ук мисаллар азагындагы “Эзләү битенә кайту” боерыгын да кулланырга мөмкин:

The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area displays a list of articles, each starting with a play button icon. The first article is titled "Шулай ук телевизордан яңгыраган куркыныч хәбәрләргә йөрәккә якин алмаска, үзең берәр мавыктыргыч шөгьль табарга, кызыклы **kitap** укырга, күбрәк жыләк-жимеш һәм шоколад ашарга кушалар." and includes a citation: "(чыганак: "Татарстан яшьләре" газетасы (web-сайт)) [Текстны табу](#)". The second article is titled "Күрәсең, безнең халыкка **kitap** байлыгы түгел, ә башка байлык кыйммәтрәк, чөнки бик кыйммәт, шундый затлы китаплар ун сум, шуның кадәр төшерделәр **kitap** бәясен, китапның бәясе төшәргә тиеш түгел, чөнки ул бик күп кешенең хезмәте." and includes a citation: "(чыганак: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)". The third article is titled "**Kitap** бизәү өлкәсендә «Жайдак» үзен тиз таныта." and includes a citation: "(чыганак: Максим Глухов. Эзләнүләр жимеше)". The fourth article is titled "Башка кунакларны көткәндә, **kitap** сүзләрен сөйләләр иде." and includes a citation: "(чыганак: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)". The fifth article is titled "1) Дүрт йөзгә якин фотосурәт төшәргән фотографларның яшен һәм елъязманьң Калининградта чыгарылуын да исәпкә алсаң, **kitap** шактый гына бәяләгә охшап тора." and includes a citation: "(чыганак: Гульшат Галиуллина. Татар теле: лексикология (таблицалар, анализ үрнәкләре, күнегүләр, сүзлекчә). КАЗАН - 2007)".

Navigation links at the bottom of the article list include: < Артка | Беренче биткә | Алга >. Below these links, there are two buttons: [Өскә] and ["Эзләү" битенә кайту], where the second button is highlighted with a red border.

At the bottom of the page, the text "Эксперименталь фонетика лабораториясе" is visible.

Татар хәрефләрен ничек язарга белмәсәгез!

Өгәр компьютерыгызда татар алфавиты булмаса, Корпусның эзләү битендә сез кирәкле сүзне виртуаль клавиатура ярдәмендә яза аласыз. Моның өчен “**Табу!**” тәймәсеннән сулда урнашкан кечкенә клавиатура рәсеменә чиртегез:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG баш биткә

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

Татар теленең язма корпусында эзләү

Эзләү төрөн сайлап, кирәкле сүзне языгыз:

Табу!

☉ Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)

○ Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)

○ Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *лөргә**)

Эксперименталь фонетика лабораториясе

Шуннан соң алдыгызда клавиатура тәрәзәсе ачылачак. Аның тиешле төймәләренә басып сез татар хәрефләреннән торган сүзне җыя алачаксыз:

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". On the left, there is a navigation menu with options: "ТАТ РУС ENG", "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустаң файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтөгә керү: tatcorpus@gmail.com", and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". The main content area has the heading "Татар теленең язма корпусында эзләү" and a search prompt "Эзләү төрен сайлап, кирәкле сүзне языгыз:". Below this is a search input field and a "Табу!" button. A list of search options is provided: "Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)", "Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)", and "Шаблон буенча эзләү (3 һәм артырак хәрефтән торган сүз башы дүм эзләгү кирәктелә, мәсәлән)". A "Virtual Keyboard - Mozilla Firefox" window is overlaid on the page, showing a Tatar keyboard layout with keys for letters like 'ә, һ, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, -, =, \, /, Bk.sp', 'Tab, й, ц, ө, у, к, е, н, г, ш, щ, э, з, х, ь, ү', 'CapsLock, ф, ы, в, а, п, р, о, л, д, ң, э, Enter', 'Shift, я, ч, с, м, и, т, ь, ж, <, >, ., ', Shift, Del', and 'Ctrl, Alt, Alt, Ctrl'. The keyboard version is noted as "VirtualKeyboard 3.7.2.797" and the language is set to "Tatar".

Эзләү нәтижеләренә туры сылканы алу

Эзләү нәтижеләренә туры сылқа буларак браузерның адрес юлындагы мәгълүматны куллану дәрәс түгел. Корпус сайты фреймлы структурага корылган булу сәбәпле, адрес юлындагы мәгълүмат бары тик сайтның баш битенә генә сылтама ясей. Нәкъ менә эзләү нәтижеләренә туры сылқа алып өчен әлеге битнең иң өске өлешендәге «**Туры сылқа**» кнопкасына басарга кирәк:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG ["Эзләү" битенә кайту] [Жәмләләргә күчү] [**Туры сылқа**] баш биткә

Баш бит
Корпуса эзләү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

СҮЗ: **китан** ▶

ТАБЫЛУ САНЫ: 31866

УҢ КОНТЕКСТ: нәшриятында (1409; 19717.9), нәшрияты (1026; 14194.5), булып (871; 3345.21), авторы (869; 8785.8), итеп (787; 3081.54), " (628; 620.15), та (555; 2390.1), белән (476; 140.15), дөнья (407; 2037.03), » (392; 435.78), укырга (382; 2465.02), нәшриятының (379; 5576.17), укучылар (305; 1878.74), укып (304; 1943.77), - (303; 43.75), чыгару (262; 1630.57), күргәзмәсе (254; 2204.72), уку (252; 1002.89), нәшер (237; 1964.75), бүләк (204; 972.04), бастырып (196; 1450.99), турында (184; 205.99), бастыру (184; 1679.97), басылып (173; 1192.09), чыгарган (163; 1126.32), рөвешендә (160; 1110.37), чыгарырга (160; 1137.4), басу (159; 1292.96), укучы (157; 831.22), укый (150; 766.25), нәшриятына (149; 2191.66), алып (144; 133.16), " (143; 12.47), язган (142; 690.19), язып (142; 709.43), сөючеләр (134; 1248.08), — (132; 20.56), язарга (131; 853.71), укымый (127; 1443.4), чыкты (126; 399.29), кибетә (126; 1205.3), нәшрияты (125; 2116.87), бар (123; 59.78), чыгара (117; 659.77), чыгарды (113; 743.19), яза (112; 496.18), язу (105; 623.5), укуга (101; 951.18), язарлык (100; 1228.2), сәүдәсе (99; 1153.05), кибетләрендә (99; 1098.99), сатып (98; 307.17), - (97; 4.65), күргәзмәләре (96; 1068.96), кибетендә (94; 1000.95), төкъдим (92; 169.96), кибетләре (91; 973.71), буларак (81; 153.79), кибетенә (80; 861.09), сүзе (74; 339.54), чыгарып (74; 347.04), фонды (70; 280.96), кына (69; 99.87), сату (65; 288.18), киштәсе (65; 920.46), килеп (63; 36.21), укыган (63; 286.33), өстендә (61; 220.12), авторлары (56; 416.63), киштәләрендә (56; 616.84), кебек (55; 14.85), киштәләре (55; 669.43), төкъдир (53; 303.83), аша (50; 64.75), теле (50; 89.08), тотып (50; 137.34), кибетләренә (50; 582.79), алырга (49; 86.31), нәшрияте (49; 637.86), чыккан (48; 66.4), укучыларны (48; 277.7), тарату (48; 315.49), алам (46; 201.67), укыйм (46; 313.4), күргәзмә-ярминкәсе (45; 516.95), нәшр (43; 445.79), укуны (42; 332.42), яки (41; 34.96), укучыларга (40; 194.98), нәшриятлары (40; 490.78), укучыларның (39; 169.41), язды (38; 206.85), дөньясына (38; 214.69), нәшриятендә (38; 487.07), сөюче (38; 309.94), битләре (38; 318.18), киштәсендә (37; 467.97), битләренә (36; 322.77), бастыра (36; 229.48), палатасы (35; 151.32) Барысын да карау

СУЛ КОНТЕКСТ: %^% (2870; 265.87), татарстан (2781; 16768.4), , (2522; 144.94), бу (1203; 2840.53), аерым (771; 4948.61), " (709; 742.16), бер (680; 1378.6), _NUMBER_ (660; 66.09), дигән (659; 2501.77), турында (507; 1398.4), « (365; 325.43), әлеге (362; 926.18), яна (359; 865.88), исемле (350; 2094.36)

corpus2.tatfolk.ru:8083/corpus/search/find.php?keyword=китан&keyword2=сатып&pos=right&loglikelihood=1&lang=tt

Ачылачак тәрәзәдә сезгә кирәкле сыылка урнашкан булыр:

ТАТАР ТЕЛЕНЕН ЯЗМА КОРПУСЫ

ТАТ РУС ENG ["Эзләү" битенә кайту] [Жөмлөләргә күчү] [Туры сыылка] баш биткә

Баш бит
Корпуста эзләү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

СУЗ: *китап* ▶

ТАБЫЛУ САНЫ: 31866

УҢ КОНТЕ
авторы (869; 14194.5), булып (871; 3345.21),
дөнья (407; 2390.1), белән (476; 140.15),
укучылар нәшриятының (379; 5576.17),
күргәзмәсе чыгару (262; 1630.57),
бастырып (964.75), бүлөк (204; 972.04),
чыгарган (79.97), басылып (173; 1192.09),
укучы (157; 1137.4), басу (159; 1292.96),
укучы (157; 144; 133.16), " (143; 12.47),
язган (142; 690.19), язып (142; 709.43), сөючеләр (134; 1248.08), — (132; 20.56), язарга (131; 853.71),
укумый (127; 1443.4), чыкты (126; 399.29), кибете (126; 1205.3), нәшрияты (125; 2116.87), бар (123; 59.78),
чыгара (117; 659.77), чыгарды (113; 743.19), яза (112; 496.18), язу (105; 623.5), укуга (101; 951.18),
язарлык (100; 1228.2), сөүдәсе (99; 1153.05), кибетләрендә (99; 1098.99), сатып (98; 307.17), - (97; 4.65),
күргәзмәләре (96; 1068.96), кибетендә (94; 1000.95), төкъдим (92; 169.96), кибетләре (91; 973.71),
буларак (81; 153.79), кибетенә (80; 861.09), сүзе (74; 339.54), чыгарып (74; 347.04), фонды (70; 280.96),
кына (69; 99.87), сату (65; 288.18), киштәсе (65; 920.46), килеп (63; 36.21), укыган (63; 286.33),
өстендә (61; 220.12), авторлары (56; 416.63), киштәләрендә (56; 616.84), кебек (55; 14.85),
киштәләре (55; 669.43), төкъдир (53; 303.83), аша (50; 64.75), төле (50; 89.08), тотып (50; 137.34),
кибетләренә (50; 582.79), алырга (49; 86.31), нәшрияте (49; 637.86), чыккан (48; 66.4),
укучыларны (48; 277.7), тарату (48; 315.49), алам (46; 201.67), укыйм (46; 313.4), күргәзмә-
ярминкәсе (45; 516.95), нәшр (43; 445.79), укуны (42; 332.42), яки (41; 34.96), укучыларга (40; 194.98),
нәшриятлары (40; 490.78), укучыларның (39; 169.41), язды (38; 206.85), дөньясына (38; 214.69),
нәшриятендә (38; 487.07), сөюче (38; 309.94), битләре (38; 318.18), киштәсендә (37; 467.97),
битләренә (36; 322.77), бастыра (36; 229.48), палатасы (35; 151.32) Барысын да карау

СУЛ КОНТЕКСТ: %^% (2870; 265.87), татарстан (2781; 16768.4), , (2522; 144.94), бу (1203; 2840.53),
аерым (771; 4948.61), " (709; 742.16), бер (680; 1378.6), NUMBER (660; 66.09), дигән (659; 2501.77),
турында (502; 1398.4) « (365; 325.43) алегә (362; 926.18) яна (359; 865.88) исемле (350; 2094.36)

Эксперименталь фонетика лабораториясе

The page at <http://corpus2.tatfolk.ru:8083> says:

Туры сыылка

[/corpus/search/find.php?keyword=%D0%BA%D0%B8%D1%82%D0%B0](http://corpus/search/find.php?keyword=%D0%BA%D0%B8%D1%82%D0%B0)

Корпуста аерым сүзформа һәм аның контекстына бәйле статистик мәгълүматны табу

Тәрәзәнең текст аланчыгына кирәкле сүзне языгыз (мисал өчен: **китапның**).

Игътибар итеgez! Берез түбәндәрәк “**Коллокацияләр (статистик) корпусында сүзформа эзләү...**” дип башланган юл билгеләнгән булырга тиеш. Шуннан сул кнопка белән “**Табу!**” тәймәсенә басыгыз.

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area is titled "Татар теленең язма корпусында эзләү" (Search in the Tatar written corpus). Below the title, there is a search prompt: "Эзләү төрен сайлап, кирәкле сүзне языгыз:" (Choose a search type, write the word you need). A search input field contains the word "китапның" (book's), and a "Табу!" (Find!) button is next to it. Below the search field, there are three radio button options for search types:

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *лөргә**)

On the left side of the page, there is a sidebar with navigation links: "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустан файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында". Below these are contact information: "Элемтөгә керү: tatcorpus@gmail.com" and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". At the bottom of the page, it says "Эксперименталь фонетика лабораториясе".

Аерым тәрәзәдә эзләү нәтижәләре ачылачак. Монда башта эзләнәсе сүзформа үзе һәм аның Корпуста очрау ешлыгы күрсәтелә. Шуннан соң өч бүлек бирелә:

- **Уң контекст**, ягъни табылган сүзформа артыннан килүче сүзләрнең очрау ешлыклары буенча тәртипкә салынган жыелмасы. Жыелмадагы һәр сүзнең уң ягында жәяләр эчендә ике сан күрсәтелә:

1. **саргылт төстәге** сан Корпуста бу сүзтәзмәнең (**китапның** + уң контексттагы кайсыдыр сүз, мисал өчен: **эчтәлеген**) ничә тапкыр очравын күрсәтә. Бу очракта, күрәбез ки, “**китапның эчтәлеген**” корпуста 26 тапкыр очрый.

2. **яшел төскә** буялган сан **логарифмик дәрәслеккә тартымлык күрсәткечен** (log-likelihood) тәшкит итә. Бу төшенчә турында тулырак мәгълүматны Тед Даннингның (Ted Danning) шәхси блогында түбәндәге адрес буенча табарга мөмкин:

<http://tdunning.blogspot.ru/2008/03/surprise-and-coincidence.html>

The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The search term is "СҮЗ: китапның" and the frequency is "ТАБЫЛУ САНЫ: 3120". The "УҢ КОНТЕКСТ" section lists various words with their frequencies, with "эчтәлеген (26; 285.34)" highlighted in a red box. The "СҮЛ КОНТЕКСТ" section at the bottom shows the word "бу" with a frequency of 305 and a log-likelihood of 1270.4. The footer mentions "Эксперименталь фонетика лабораториясе".

- **Сул контекст** - эзләнелә торган сүзформаның алдыннан килүче сүзләр. Жәяләр эчендә урнашкан саннарның мәгънәсе уң контексттагы саннарның мәгънәсенә туры килә.
- **Семантик контекст** - бу сүзләр эзләнелә торган сүзформа белән бер жәмләдә очрыйлар һәм, шулай итеп, аның белән билгеле бер дәрәжәдә тематик (семантик) планда якынлык тәзиләр. Алар бу исемлектә ешлыклары кимүгә мөнәсәбәттә тәртипкә салынган.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

Баш бит
 Корпуста эзләү
 Статистик мәгълүмат
 Корпус тасвирламасы
 Корпустан файдалану
 Басылган мәкаләләр
 Рәхмәт сүзе
 Проект турында

Элемтөгә керү:
 tatcorpus@gmail.com

Авторлар:
 Сайхунов М.Р.
 Ибраһимов Т.И.

СУЛ КОНТЕКСТ: %~% (903; 1453.07), бу (305; 1270.4), , (224; 4.69), изге (150; 1333.51), әлегә (119; 576.55), ул (69; 103.51), дигән (47; 150.31), исемле (30; 171.36), яңа (27; 52.82), һәр (24; 71.81), шул (24; 33.4), бер (22; 9.61), чыккан (19; 72.8), ук (18; 19.23), булчак (17; 52.69), турындагы (16; 65.46), ө (16; 8.43), аталган (12; 62.09), шушы (12; 21.82), беренче (12; 10.21), икенче (12; 22.34), торган (11; 8.26), өһле (11; 128.83), берничә (11; 26.21), өмма (11; 15.16), язылган (10; 37.13), кулязма (10; 93.08), күргән (9; 39.25), багышланган (9; 26.72), автор (9; 46.24), басма (9; 60.61), өч (9; 17.26), борынгы (8; 31.92), ителгән (7; 16.58), мондый (7; 9.98), тик (7; 7.75), чөнки (7; 8.42), миңа (7; 12.19), исемнәре (6; 33.04), басылган (6; 32.22), тарихи (6; 17.88), алган (6; 7.74), алынган (6; 14.72), кадерле (6; 27.01), шигырьләр (5; 20.66), торучы (5; 20.74), томлы (5; 51.84), кызыл (5; 13.67), килчөктә (5; 13.1), текст (5; 34.75), томлык (4; 34.99), биографик-публицистик (4; 86.32), исемләнгән (4; 17.56), фәнни (4; 9.3), хәтта (4; 4.64), исеме (4; 8.84), андый (4; 9.43), кызыксыну (3; 11.72), кыскача (3; 16), битлек (3; 22.31), бүленгән (3; 15.88), тормышында (3; 12.86), мөшһүр (3; 12.91), тышлы (3; 27.7), уникаль (3; 15.58), калын (3; 15.57), ошбу (3; 16.82), данә (3; 16.46), ушбу (3; 29.99), аңлатмалар (3; 24.08), электрон (3; 7.87), күңелле (3; 6.62), яза (3; 6.16), * (3; 5.43), әлегә (3; 4.41), жиде (3; 8.44), һәрбер (3; 8.87), теләгән (3; 9.18), гадәти (3; 10.6), моны (3; 4.01), әсәрләрдә (3; 23.57), ләп (3; 21.02), чыгарылган (3; 11.69), кирәкле (3; 7.54), битле (2; 14.27), мөхәммәтшиннар (2; 27.08), шөжәрәсе (2; 16.98), кәшифөгә (2; 34.78), битенә (2; 12.49), жәмәгәтчеләгендә (2; 20.1), исемдәге (2; 12.57), көйле (2; 15.02), авторлар (2; 11.45), габидулла (2; 24.94), жавапны (2; 11.57), хикмәтле (2; 13), бөлкөм (2; 10.89), бар—алары (2; 43.16), әзерләнүче (2; 17.86), фикерләргә (2; 12.84) Барысын да кара

СЕМАНТИК КОНТЕКСТ: . (2934; 244.58), , (1741; 50.28), һәм (713; 187.99), _NUMBER_ (586; 39.54), бу (536; 338.75), да (431; 23.99), белән (417; 9.5), « (378; 334.19), » (353; 313.95), - (336; 24.63), дө (330; 17.62), " (318; 24.2), " (294; 22.12), ул (294; 24.14), дип (281; 39.63), : (260; 73.1), бер (254; 65.54), ((228; 142.83),) (223; 128.39), беренче (215; 249.58), турында (209; 179.09), татар (201; 116.14), авторы (201; 1294.23), әлегә (190; 179.29), дигән (187; 193.42), аның (171; 33.32), изге (164; 623.39), - (149; 19.47), китап (145; 527.08), шул (144; 39.24), ук (137; 39.36), икенче (137; 200.95), ө (136; 9.3)

Эксперименталь фонетика лабораториясе

Контекстлар бүлегеннән түбәндәрәк файдаланучыларга “Мисаллар” жыелмасы тәкъдим ителә.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Г.И.

МИСАЛЛАР:

▶ Әлеге уңайдан Бөтендөнья татар конгрессы Башкарма комитеты рәисе урынбасары Ренат Вәлиуллин **китапның** тиздән сатуга чыгачагын, шулай ук республикадан читтә дә таратылачагын белдерде.
(чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлығы (web-сайт)) [Текстны табу](#)

▶ Шуңа күрә дә китабымның изге максаты – жыйналган мәгълүматларны, тарихи чыганақлар белән тулыландырып, халкымның үзенә кайтару». 3 мең данә тираж белән нәшер ителгән **китапның** тышлығын рәссам Рәисә Сәйфуллина бизәгән.
(чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлығы (web-сайт)) [Текстны табу](#)

▶ Дүрт тематик кискәккә бүленгән **китапның** эпиграфлары ук игътибарны туплый, кызыксынуны көчәйтә.
(чыганақ: РИФӘ РАХМАН. Татар телендә 100 сочинение)

▶ **Китапның** икенче басмасы алдагы басмада жиберелгән күп кенә житешсезлекләрдән, авыр сүзләрдән азат ителгән.
(чыганақ: "Фән һәм тел" журналы (№27))

▶ Профессор Г. Ф. Саттаров “Казанчы, Сабанчы, Уракчы, Урманчы, Ындырчы, Буранчы, Туйчы, Ямгурчы һ. б. — борыңгы төрки татар кешесе исемнәре” дип яза алда телгә алынган хезмәтендә (85нче бит.) “Башкортстанда феодализм һәм капитализм тарихыннан” дигән **китапның** (Уфа, 1971 ел) 267нче битендә “Слободские татаровья без хлебно́го жалованья Итку́л Бехтеми́ров, Сабанча Чилпанов...” дигән юлларны укыйбыз.
(чыганақ: "Кызыл таң" газетасы (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Монда тиешле сүзформаны үз эченә алган яқынча 50 жәмлә китерелә. Уңайлык өчен эзләнүче сүзформа жәмләләрдә кызыл төскә буяла.

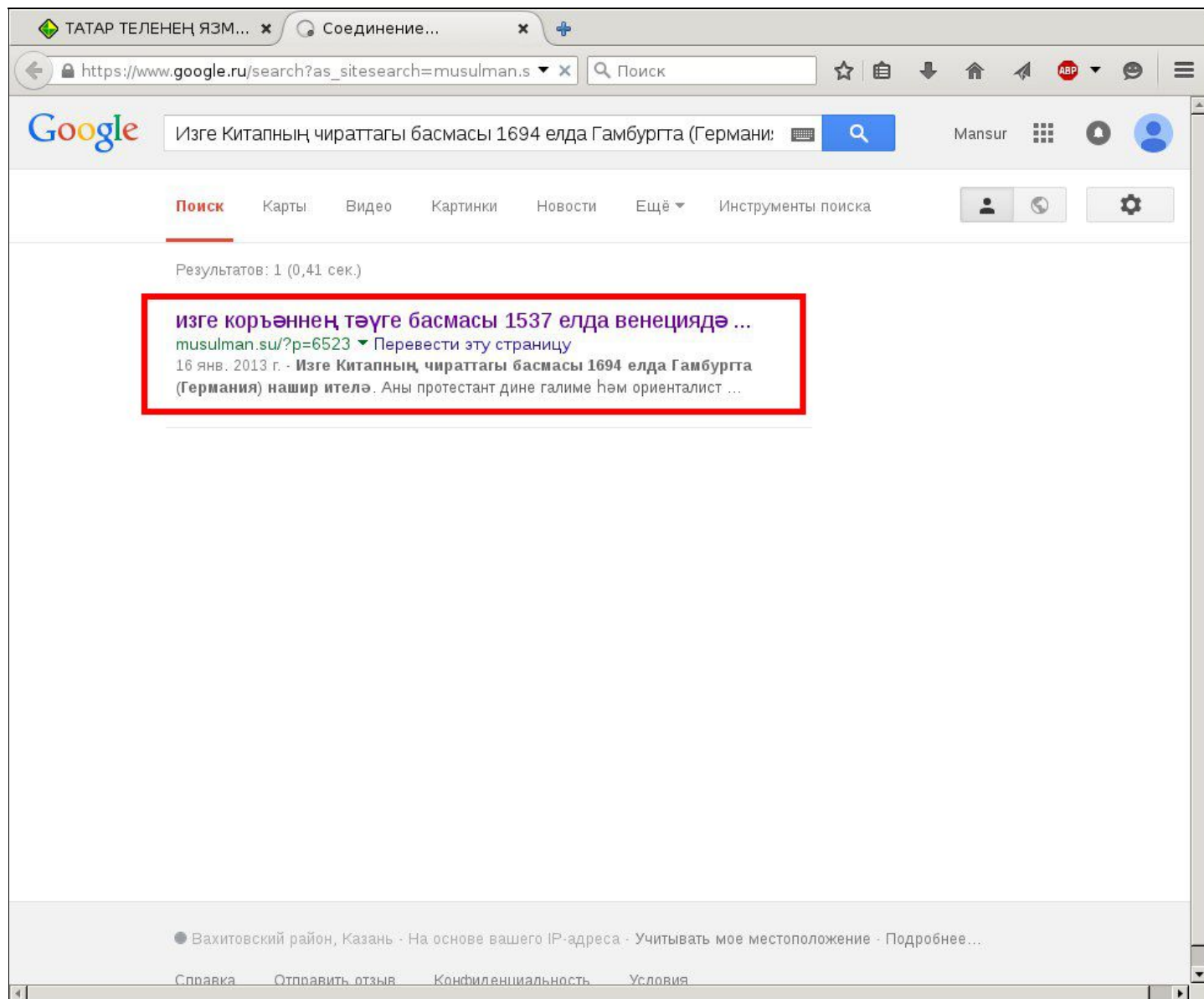
Мисалларда һәр жөмлэдән соң кечкенә шриффт белән ул жөмләнең чыганагы, ягъни аның кайдан алынган булуы әйтелә.

Чыганак веб-ресурс булган очракта (аларның барысы да “*web-caйт*” дип тамгаланган) “Текстны табу” дигән искөрмә күрсәтелә.

The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". There are navigation links for "ТАТ", "РУС", and "ENG". A sidebar on the left contains links for "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустаң файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", and "Проект турында". Below the sidebar, there is contact information: "Элемтөгә керү: tatcorpus@gmail.com" and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". The main content area displays a list of articles, each with a play button icon, a title, and source information. The second article is highlighted with a red box: "Изге **Китапның** чираттагы басмасы 1694 елда Гамбургта (Германия) нашир ителә. (чыганак: "Ислам - Татарлар һәм мәселманнар" (web-caйт)) [Текстны табу](#)". Other articles include: "Хәлбуки, табыг итгерү ноктасыннан **китапның** озын булуы матлуб түгелдер. (чыганак: "Эхо веков" журналы)", "Өйгә эш катлаулашты: әкиятне укып, үзеңә ошаган фрагментын сүрәткә төшерергә, һәм **китапның** ни хақында булуын бер жөмлә белән генә язып куярга кирәк. (чыганак: "Гыймрановларның сайты" (web-caйт)) [Текстны табу](#)", "**Китапның** әһәмиятен якташыбыз, тарих фәннәре кандидаты, Казан федераль (педагогика) университеты профессоры, ТР Мәгълүматлаштыру академиясенең действительный әгъзасы Иршат Гафаров һәм чуваш төбәкләрен өйрәнү союзының мактау-лы рәисе Виталий Станьял югары бәяләделәр. (чыганак: "Туган як" газетасы (web-caйт)) [Текстны табу](#)", "Сагыну" дигән **китапның** кыйбласы менә шул тарафта. (чыганак: "Intertat.ru" электрон газетасы (web-caйт)) [Текстны табу](#)", "Күпме хезмәт, чыгым тотып әзерләнган әлеге **китапның** тузан жыеп ятуы шагыйрәне генә түгел, каләмдәшләрен дә борчуга салган һәм алар ярдәм сорап, УР Милли сәясәт министрлыгының милли эшләр буенча баш белгече Айрат Гайнетдиновка мөрәжәгать ителәр. (чыганак: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-caйт)) [Текстны табу](#)". At the bottom of the page, it says "Эксперименталь фонетика лабораториясе".

Мисал матур әдәбиятка караган тексттан алынса, “Автор хокукын саклау” кагыйдәләренә бәйләнештә текстны табу мөмкинлеге чикләнә.

“Текстны табу” сылтамасына чиртү нәтижәсендә корпустан файдаланучы кеше Google-ның эзләү системасына юлыга.



Күп очракта искәрмәнең беренчесе үк кирәкле жөмләне эченә алган текстның битен күрсәтәчәк.

ТАТАР ТЕЛЕНЕҢ ЯЗМ... x Изге Китапның чир... x ИЗГЕ КОРЪӘННЕҢ ТӘҮГ... x

musulman.su/?p=6523

ИЗГЕ КОРЪӘННЕҢ ТӘҮГЕ БАСМАСЫ 1537 ЕЛДА ВЕНЕЦИЯДӘ ДӨНЬЯ КҮРҮЕ ТУРЫНДА ИШЕТКӘН ИДЕГЕЗМЕ?

ГАЛӘМ ADMIN 16.01.2013 12:36 0 ФИКЕР

4501.jpg

Ислам тарихында Каләм Шерифнең ташка баскан беренче үрнәкләре озак еллар бие Европада билгеле булып та, Шәрәк илләренә үтеп кермәве сәерерәк тоелыр, әмма ләкин бу тарихи фәкт XIX гасырның күренекле мәстәшиге (Шәрәкне әйрәнүче галим) М. Казымбак моны мәселман дин белгечләренең изге текстның алыш-биреш әйберенә әйләнүенән шикләнүдән килә, дип аңлата.

Бүгенге көндә Коръәннең гарәп телдәге беренче басмасы дип Италиядә нашир ителгән нәсхәләре санала. Шуларның иң тәүгесе 1530 елда Венециядә дөнья күргән. Мәгар беренче басмада күпсанлы хаталар табылганлыктан, наширләр бөтен тиражны жәеп алып яндырырга мәжбүр булалар һәм аны тәзәтеп 1537 елда бастыралар.

Изге Китапның чираттагы басмасы 1694 елда Гамбургта (Германия) нашир ителә. Аны протестант дине галиме һәм ориенталист Абрахам Хинкельман әзерли.

Тагы дүрт елдан (1698) Коръәннең гарәп телдәге тексты белән аны латин телдәге тәржәмәсе Падүя (Италия) каласында басыла. Әлеге басманың бер нәсхәсе Казан Федераль Университетының сирәк китаплар саклана торган мирасханәсендә бар.

Бары тик 1717 елда гына Госманлы империясендә Стамбулның баш мөфтиә Абдулла-әфәнде карары белән һәм солтан Әхмәд-хан манифесты буенча Коръәнне бастыру жайга салына.

Издегән текст: Изге Китапның чираттагы басмасы 1694 елда Гамбургта (Германия) нашир ителә.

Подсветить все Учетьь регистр 1-е из 1 совпадения

Достигнут низ страницы, продолжено сверху

Эзләү

Эзләргә

РАДИО

ВИДЕО

Жомга вәгазе - Айдар хәзрәт Карибуллин

Вәгазь — Зәфәр хәзрәт Галиуллин

Апанай мәчетендә жомга вәгазе —

Әмма жөмлә монда төс белән аерып күрсәтелми. Мондый очракта шәхес аны веб-браузерның эзләү функциясен кулланып (клавиатурада берьюлы **Ctrl** һәм **F** тәймәләрен басып), я булмаса текстны карап чыгып таба ала.

Мисал жөмлөләр жыелмасының астында файдаланучыга башка битлэргә чыгу мөмкинлеге бирелгән. Әгәр шәхес сул кнопка белән “Алга” дигән сүзгә чиртсә, эзлөү системасы файдаланучыга яңа 50 жөмлә тәкъдим итә. Хезмәткәрләр шушы мөмкинлектән кабат-кабат файдаланып үзләренә кирәкле сүзформа ярдәмендә төзелгән Корпустагы барлык жөмлөләр белән танышып чыга ала.

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". On the left side, there is a navigation menu with options: "ТАТ РУС ENG", "Баш бит", "Корпуста эзлөү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустан файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтәгә керү: tatcorpus@gmail.com", and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". The main content area features a list of articles, each starting with a play button icon and the word "Китапның" in red. The articles are:

- Китапның** авторы — алдынгы фикерле гарәп (Мисыр) язучысы (чыгышы белән көрд) Касыйм Әмин (1865—1909).
(чыганақ: "Габдулла Тукайга багышланган Интернет-портал" (web-сайт)) [Текстны табу](#)
- Һәм шушы **китапның** дөнья күрүендә Сезнең ярдәм итүегезне сорыйбыз.
(чыганақ: "Туганайлар" газетасы (web-сайт)) [Текстны табу](#)
- Китапның** беренче өлеше ел фасыллары белән бәйлә горөф-гадәтләрне үз эченә алса, икенче өлешендә гаилә традицияләре каләмгә алынган.
(чыганақ: "Белем.ру" порталы (web-сайт)) [Текстны табу](#)
- Китапның** беренче кисәгендә 30-35 документ файдаланылган.
(чыганақ: РИЗАЭДДИН ФӘХРЕДДИН: МИРАСЫ ҺӘМ ХӘЗЕРГЕ ЗАМАН: Мәкаләләр әҗыентыгы. Казан, 16 ноябрь 1999 ел. - Казань, 2003.)
- Китапның** кулъязма вариантларында һәм беренче (1847) басмасында Т. Ялчыголның үзе иҗат иткән байтак шигъри юллар бар.
(чыганақ: "Татар әдипләре" порталы (web-сайт)) [Текстны табу](#)
- Өч **китапның** да үзәген кеше гомере, күңелдәге кичерешләр һәм дөнья турында уйланулар тәшкил иткәнлектән алар бер-берсен тулыландыралар.
(чыганақ: "Татар әдипләре" порталы (web-сайт)) [Текстны табу](#)
- Китапның** төп өлешен архив документлары, ягъни ревизия материаллары алып тора да.
(чыганақ: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)

At the bottom of the page, there is a navigation bar with three buttons: "< Артка", "Беренче биткә", and "Алга >". The "Алга >" button is highlighted with a red box. Below the navigation bar, the text "Эксперименталь фонетика лабораториясе" is visible.

Корпуста аерым лексеманы табу

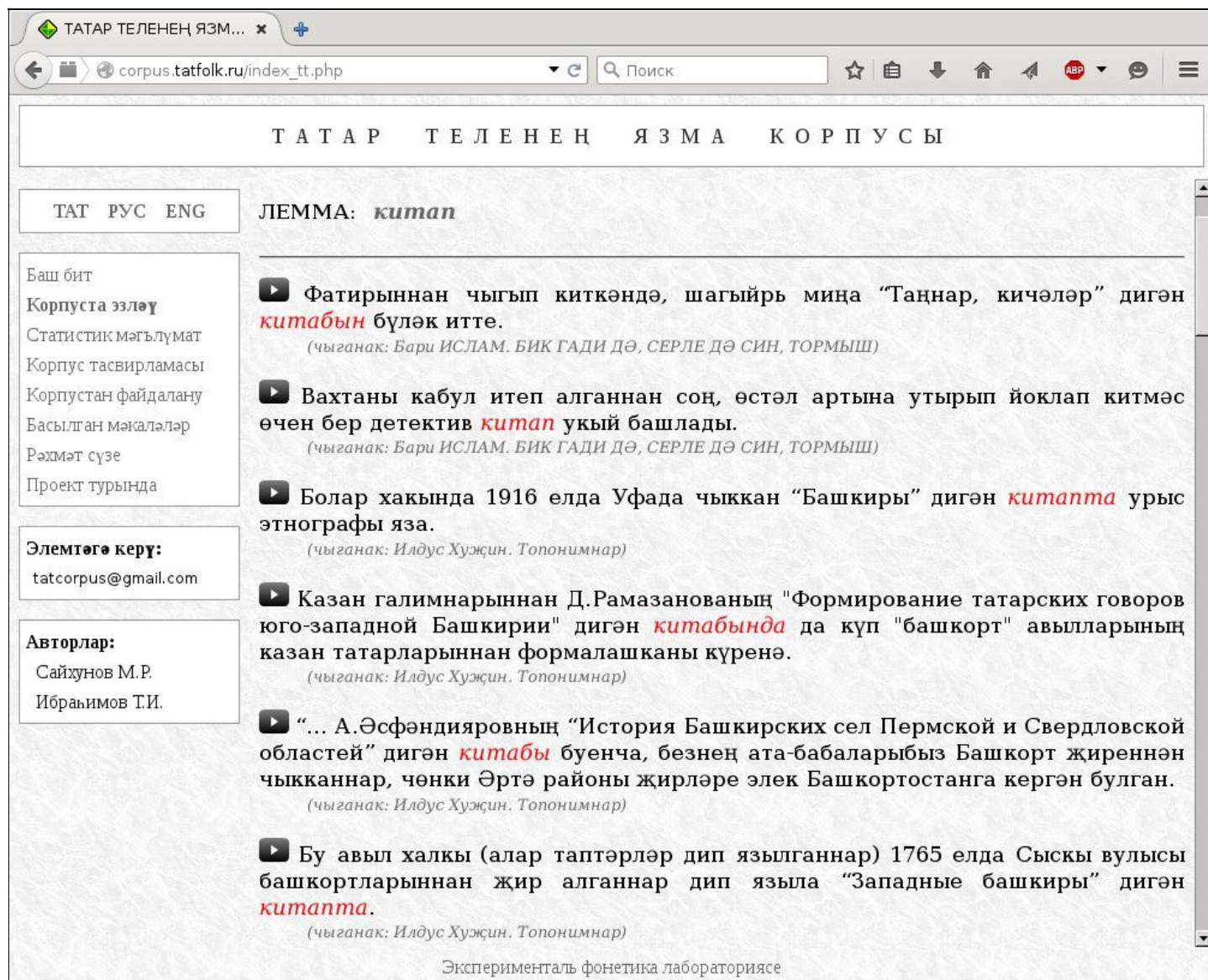
Язма корпус бүгенге көндә тулысынча морфологик яктан тамгаланган. Өгәр файдаланучыга аерым леммага караган сүзформалар кERGән жәмләләрне табу кирәк булса, ул текст аланчыгына лемманы (мәсәлән, **тормыш**, **тутыр**, **матур**) язарга һәм “**Морфологик тамгаланган корпуста лемма эзләү...**” юлын билгеләргә тиеш була.

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area displays the search results for the word "китап". The search bar contains "китап" and a "Табу!" button. Below the search bar, there are three radio button options for filtering the results:

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны**, **авылларга**, **килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап**, **авыл**, **кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап***, **авыл***, ***гез**, ***лөргә**)

The second option is selected and highlighted with a red box. The footer of the page reads "Эксперименталь фонетика лабораториясе".

“**Табу!**” боерыгына чиртү белән файдаланучыга яңа бит ачыла. Бу биттәге жәмләләрдә аланчыкка язылган лемма төркеменә кергән формалар күзәтеләчәк.



The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The search results are for the lemma "kitap".

Navigation links: TAT RUS ENG

Left sidebar menu:

- Баш бит
- Корпуста эзләү
- Статистик мәгълүмат
- Корпус тасвирламасы
- Корпустаң файдалану
- Басылган мәкаләләр
- Рәхмәт сүзе
- Проект турында

Элемтөгә керү: `tatcorpus@gmail.com`

Авторлар: Сайхунов М.Р., Ибраһимов Т.И.

ЛЕММА: **kitap**

- ▶ Фатириннан чыгып киткәндә, шагыйрь миңа “Таңнар, кичәләр” дигән **китабын** бүлөк итте.
(чыганақ: Бари ИСЛАМ. БИК ГАДИ ДӘ, СЕРЛЕ ДӘ СИН, ТОРМЫШ)
- ▶ Вахтаны кабул итеп алганнан соң, өстәл артына утырып йоклап китмәс өчен бер детектив **kitap** укый башлады.
(чыганақ: Бари ИСЛАМ. БИК ГАДИ ДӘ, СЕРЛЕ ДӘ СИН, ТОРМЫШ)
- ▶ Болар хакында 1916 елда Уфада чыккан “Башкиры” дигән **kitapта** урыс этнографы яза.
(чыганақ: Илдус Хужин. Топонимнар)
- ▶ Казан галимнарыннан Д.Рамазанованың "Формирование татарских говоров юго-западной Башкирии" дигән **китабында** да күп "башкорт" авылларының казан татарларынан формалашканы күренә.
(чыганақ: Илдус Хужин. Топонимнар)
- ▶ "... А.Әсфәндияровның “История Башкирских сел Пермской и Свердловской областей” дигән **китабы** буенча, безнең ата-бабаларыбыз Башкорт жиреннән чыкканнар, чөнки Әртә районы жирләре элек Башкортостанга кергән булган.
(чыганақ: Илдус Хужин. Топонимнар)
- ▶ Бу авыл халкы (алар таптәрләр дип язылганнар) 1765 елда Сыскы вулысы башкортларынан жир алганнар дип языла “Западные башкиры” дигән **kitapта**.
(чыганақ: Илдус Хужин. Топонимнар)

Эксперименталь фонетика лабораториясе

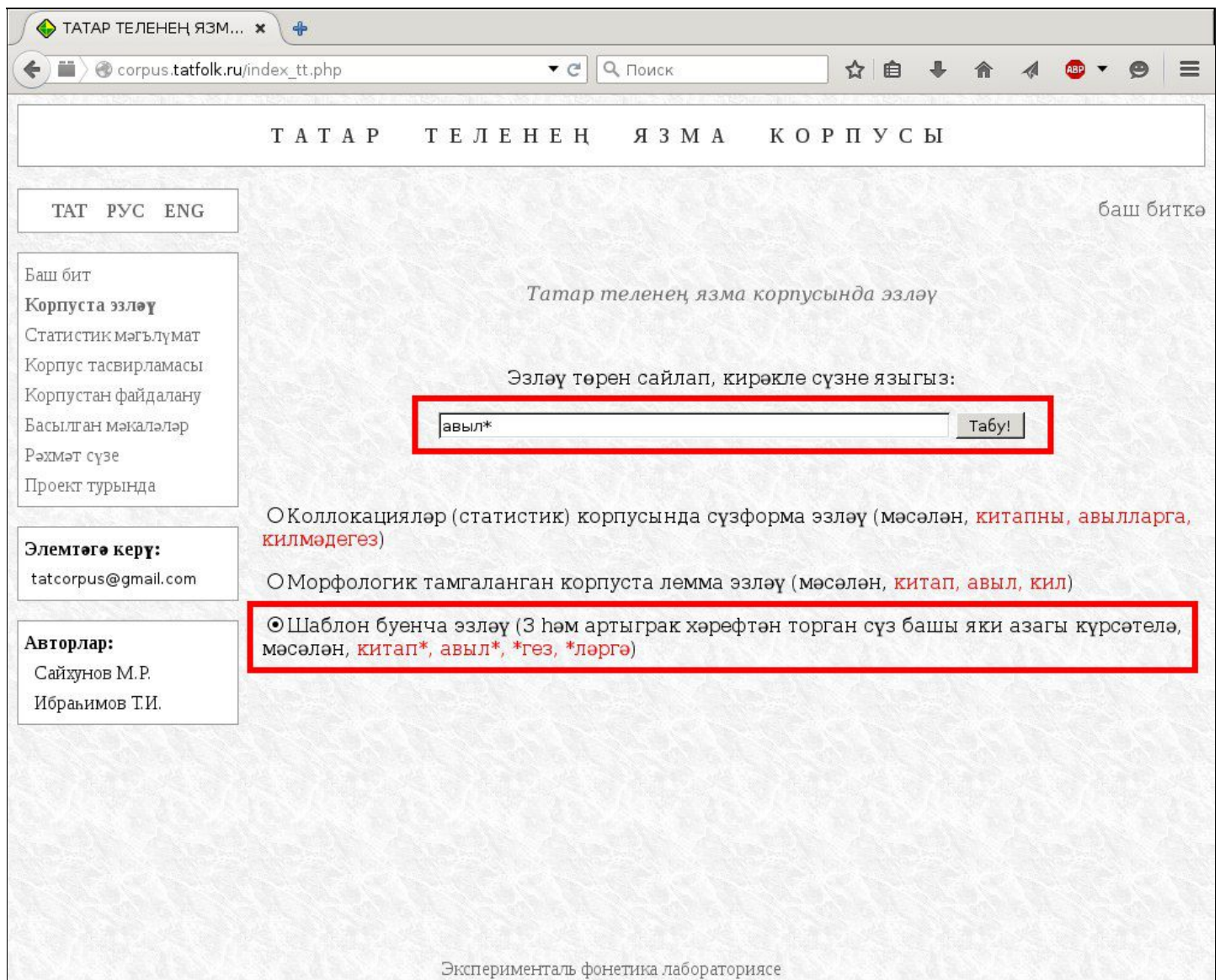
Жәмләләрнең чыганақларын күрсәтү, аларның тулы текстлары белән танышу алда күрсәтелгән кагыйдәләр нигезендә башкарыла.

Корпуста аерым префикска башланган сүзләрне табу

Корпуста бер үк хәрефләр җыелмасына башланган (префиксларга) сүзләрне табу мөмкинлеге каралган. Аның өчен эзләү битен ачып текст аланчыгына кирәкле префиксны язу һәм **“Шаблон буенча эзләү”** юлын тамгалау җитә.

Игътибар итегез! Бүгенге көндә әлеге эзләү системасы өч һәм артыграк хәрефтән торган префикслар белән генә эш итә.

Өгәр кемдер **“авыл”** сүзенә башланган сүзформалар белән танышырга теләсә, ул текст аланчыгына **“авыл*”** сүзен язарга һәм **“Табу!”** кнопкасына чиртергә тиеш була.



ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG баш биткә

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

авыл* Табу!

○ Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)

○ Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)

● Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *лөргә**)

Эксперименталь фонетика лабораториясе

Нәтижәдә файдаланучыга түбәндәге бит ачыла.

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". There are navigation links for "ТАТ", "РУС", and "ENG". The search results are for the word "авыл*".

ТАТ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

СҮЗ: авыл*

- ▶ Шушыны күрә торып **авыл** жирлеге рәисе берни эшләми, ә **авыл** тәрбияләгән Русия мөфтиләр шурасы рәисе Равил Гайнетдин килеп бер хәлебезне дә белгәне юк, ди **авыл** халкы.
(чыганақ: "Ислам - Татарлар һәм мөселманнар" (web-сайт)) [Текстны табу](#)
- ▶ Алар район үзәгендәге кибетләргә рейдлар ясыйлар, кайчакта **авыл** кибетләрендә дә булалар.
(чыганақ: "Авил таңнары" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Бу вакыйгадан соң буага «Актаныш» буасы исемен кушалар, һәм шуннан соң бу исем елгага да күчә, елга исеменнән **авыл** да «Актаныш:» исемен йөртә.
(чыганақ: "Татар әкиятләре" (web-сайт)) [Текстны табу](#)
- ▶ Метрәй **авыл** жирлеге үз эченә биш **авыл**: Метрәй (административ үзәк), Салауз Мухан, Ольгино, Иске Чакмак һәм Үрнәк **авыл**ларын ала.
(чыганақ: "Мөслим районының официал мәгълүмат сайты" (web-сайт)) [Текстны табу](#)
- ▶ Әмма **авыл** хужалыгы эшкуарлары бер-берсеннән аерым хезмәт куя.
(чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)
- ▶ **Авил** жирлекләре урамнарындагы электр энергиясенә кулланылышы хақында да ризасызлыктар бик күп.
(чыганақ: "Туган як" газетасы (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Корпуста билгеле бер хәрефләр жыелмасына тәмамланган сүзләрне табу

Корпуста бер үк хәрефләр жыелмасына тәмамланган (постфиксларга) сүзләрне табу мөмкинлеге каралган. Аның өчен эзләү битен ачып текст аланчыгына кирәкле постфиксны язу һәм “**Шаблон буенча эзләү**” юлын тамгалау җитә.

Игътибар итегез! Бүгенге көндә әлеге эзләү системасы өч һәм артыграк хәрефтән торган постфикслар белән генә эш итә.

Әгәр кемдер “**гез**” хәрефләр тезмәсенә тәмамланган сүзформалар белән танышырга теләсә, ул текст аланчыгына “***гез**” сүзен язарга һәм “**Табу!**” кнопкасына чиртергә тиеш була.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG баш биткә

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

*гез Табу!

○ Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, китапны, авылларга, килмәдегез)

○ Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, китап, авыл, кил)

○ Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, китап*, авыл*, *гез, *ләргә)

Эксперименталь фонетика лабораториясе

Нәтижәдә файдаланучыга түбәндәге бит ачыла.

The screenshot shows a web browser window with the address bar displaying 'corpus.tatfolk.ru/index_tt.php'. The page title is 'ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ'. Below the title, there are language selection buttons for 'ТАТ', 'РУС', and 'ENG', and a search bar containing the text '*гез'. The main content area displays a list of search results, each starting with a play button icon. The results include:

- Бу ханнар арасындагы көрәш Кече Мөхәммәтнең жиңүе белән төгәлләнә һәм ул Олы Урдага нигез салучы булып санала.**
(чыганак: "Татарча текстлар, программалар" (web-сайт)) [Текстны табу](#)
- Узган атнада бәйсез табиблар уздырган экспертиза, журналистны хастаханәдә тоту өчен нигез юк, дип белдергән иде.**
(чыганак: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- Чагыштырмача күптән түгел генә нигез салынган мәчетне төзүне тиз тоттылар.**
(чыганак: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)
- Европа өлгесендәге татар театрына нигез салучы бит син.**
(чыганак: Низами Р.М. Без яшибез... - Казан: Татарстан китап нәшрияты, 2011.)
- Бүген биредә "Казан заманча төргәкләү заводы"ның нигез ташлары салынды.**
(чыганак: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)
- Безгә крестьян ихтилаллары дип аңлатылган ул вакыттагы Степан Разинның бөтенроссия гражданнар сугышын, чынбарлыкта исә Алтын урданың законлы чын варислары булган Олы урда - Сарай татарлары, (казаклары дип тә әйтергә була) белән Мәскәү арасында Алтын урда империясе тәхете өчен барган иң канкоешлы, иң аяусыз, иң компромиссыз,**

Эксперименталь фонетика лабораториясе

On the left side of the page, there is a sidebar with navigation links: 'Баш бит', 'Корпуста эзләү', 'Статистик мәгълүмат', 'Корпус тасвирламасы', 'Корпустан файдалану', 'Басылган мәкаләләр', 'Рәхмәт сүзе', 'Проект турында'. Below these are contact details: 'Элемтөгә керү: tatcorpus@gmail.com' and 'Авторлар: Сайхунов М.Р., Ибраһимов Т.И.'

Корпуста кирәкле сүзтезмәне табу

Корпуслар, нигездә, түбәндәге ике юнәлештә

- телне ана тел, я булмаса чит тел буларак укыту, өйрәтүне һәм
- телне бүгенге мәгълүматлашу технологияләренә иңдерү – машина тәржемәсе, компьютерлар белән табигый телдә аралашу проблемаларын чишүдә файдалануны күздә тотып төзеләләр.

Корпусның кереш өлешендә Татар теле язма корпусының икенче төргә каравы бәян ителгән иде. Машина тәржемәсе, тел ияләренең компьютерлар белән тыгыз аралашуына җаваплы проблемаларның берсе ул – күпмәгънәлелек. Бүгенге көндә бу проблема контекстка бәйләнештә чишелә.

Югарыда Язма корпуста һәр сүзформаның сул һәм уң ягында (аның алдыннан һәм артыннан) килүче сүзләрнең күрсәтелүе әйтелде. Әмма еш кына янәшәдәге контекстлар тикшерелүче сүзформаның мәгънәсен ачыклап бетерми. Җитмәгән мәгълүматны бары тик сул яки уң контекстны тулыландырып (озынайтып) кына табарга мөмкин.

Әйттик, файдаланучының максаты “**китапның эчтәлеген**” корылмасы кергән жөмлөләрне табудан гыйбарәт ди. Моңа ирешүнең өч юлы бар:

1. Уң контекст аша. Әгәр кешене “эчтәлеген” сүзеннән соң килүче контекст кызыктырса, ул текст аланчыгына “**китапның**” сүзен язарга һәм “**Коллокацияләр (статистик) корпусында сүзформа эзләү...**” дип язылган юл очындагы кнопкага тамга куярга тиеш була. Боларны үтәгәннән соң “**Табу!**” боерыгына чирткәч, файдаланучыга яңа тәрәзә ачыла.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит
Корпуса эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхуннов М.Р.
Ибраһимов Т.И.

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

китапның Табу!

Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, китапны, авылларга, килмәдегез)

Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, китап, авыл, кил)

Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, китап*, авыл*, *гез, *лөргә)

Эксперименталь фонетика лабораториясе

Ачылган тәрәзәнең “Уң контекст” өлешендә “эчтәлеген” сүзен табып, аңа сул кнопка белән чиртергә кирәк була.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG

СҮЗ: **китапның**

ТАБЫЛУ САНЫ: 3120

УҢ КОНТЕКСТ: авторы (146; 1625.81), _NUMBER_ (144; 134.06), беренче (119; 579.32), икенче (86; 475.46), исемә (73; 567.74), төп (56; 282.71), исем (55; 453.97), бер (52; 84.03), соңгы (50; 232.22), эчтәлегә (44; 476.6), ахырында (42; 342.94), « (39; 39.47), һәр (39; 152.04), кереш (37; 400.51), “ (30; 3.95), иң (29; 59.37), авторлары (28; 299.1), тиражы (26; 323.12), **эчтәлеген (26; 285.34)**, мөхәррире (26; 226.15), исемен (22; 139.32), эчтәлегенә (21; 262.48), төзүчесе (20; 259.23), баш (17; 47.7), кадрен (16; 134.95), титул (16; 222.34), фәнни (15; 71.71), язылу (14; 117.73), әһәмияте (14; 113.48), дөньяга (13; 62.35), тышлыгы (13; 190.91), тышлыгында (13; 192.15), яңа (13; 10.95), тагын (13; 19.43), электрон (12; 62.8), кыйммәте (12; 122.25), редакторы (12; 109.51), ахырына (12; 77.81), рус (12; 40.54), тулыландырылган (11; 145.04), басылып (11; 66.24), үз (11; 4.91), эченә (11; 58.35), башында (11; 48.32), дөнья (11; 28.38), нинди (11; 21.16), берничә (11; 26.21), зурлыгы (11; 110.07), дәвам (10; 89.46), азагында (10; 69.35), ике (10; 7.43), эчендә (9; 23.93), аерым (9; 21.22), беренчесе (9; 61.76), бәясә (9; 45.17), күп (9; 4.68), тарихы (9; 38.27), үзен (9; 30.2), тышлыгын (9; 143.19), исеменә (9; 52.86), өченчә (9; 31.36), дәрәжәсен (8; 48.51), рәссамы (8; 66.54), үзенә (8; 23.84), чыгуы (8; 57.27), тышлыгына (7; 98.96), идея (7; 47.72), буеннан-буена (7; 78.26), илаһи (7; 49.64), тулы (7; 19.42), тышына (7; 80.96), максаты (7; 28.99), кулъязмасы (7; 74.51), тәржемәсен (7; 82.86), концепциясе (7; 61.05), дүртөнчә (7; 35.75), ролен (6; 34.36), алгы (6; 40.17), берсен (6; 33.5), тәүгә (6; 28.69), битләрен (6; 57.76), бәясен (6; 41.32), нигезенә (6; 48.04), урыны (6; 21.71), җаваплы (6; 23.14), нигезен (6; 43.45), авырлыгы (6; 44.28), барлык (6; 7.04), калган (6; 9.22), киңәйтелгән (5; 36.03), битләре (5; 44.65), тәржемәсе (5; 42.59), архив (5; 34.78), озакламый (5; 31.25), тышында (5; 56.4), баштагы (5; 41.35), кырына (5; 45.43), россиягә (5; 34.23), эчтәлегендә (5; 61.1), рецензенты (5; 86) Барысын да карау

СУЛ КОНТЕКСТ: %^% (903; 1453.07), бу (305; 1270.4), , (224; 4.69), изге (150; 1333.51), әлегә (119; 576.55), ул (69; 103.51), дигән (47; 150.31), исемле (30; 171.36), яңа (27; 52.82)

Эксперименталь фонетика лабораториясе

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхуннов М.Р.
Ибраһимов Т.И.

Нәтижәдә өлеге сүз кызыл төскә буяла:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

СҮЗ: **китапның**
ТАБЫЛУ САНЫ: 3120

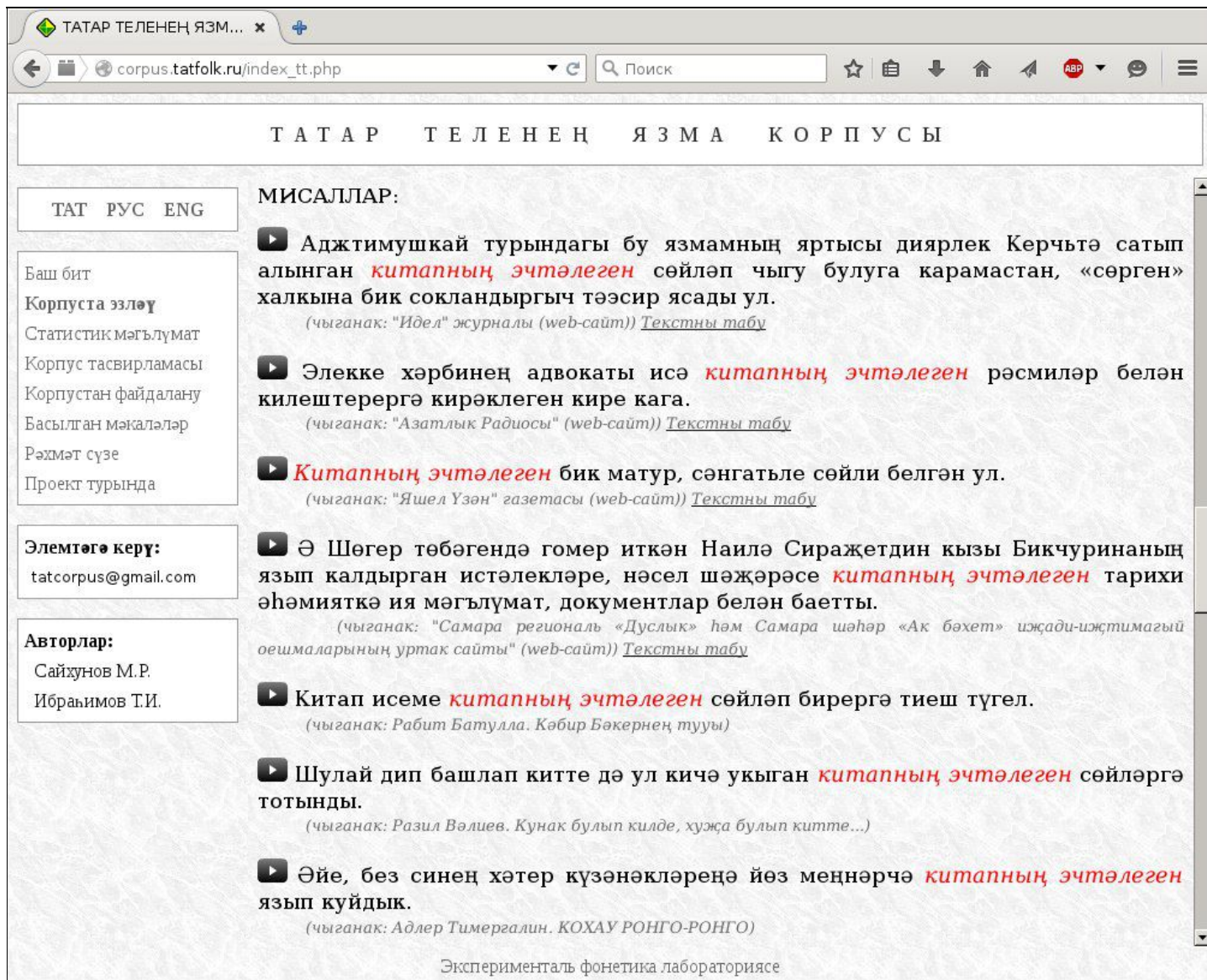
УҢ КОНТЕКСТ: авторы (146; 1625.81), _NUMBER_ (144; 134.06), беренче (119; 579.32), икенче (86; 475.46), исемә (73; 567.74), төп (56; 282.71), исем (55; 453.97), бер (52; 84.03), соңгы (50; 232.22), эчтөлөгә (44; 476.6), ахырында (42; 342.94), « (39; 39.47), һәр (39; 152.04), кереш (37; 400.51), " (30; 3.95), иң (29; 59.37), авторлары (28; 299.1), тиражы (26; 323.12), **эчтөлөгән** (26; 285.34), мөхәррире (26; 226.15), исемән (22; 139.32), эчтөлөгәнә (21; 262.48), төзүчесә (20; 259.23), баш (17; 47.7), кадрен (16; 134.95), титул (16; 222.34), фәнни (15; 71.71), язылу (14; 117.73), әһәмияте (14; 113.48), дөньяга (13; 62.35), тышлыгы (13; 190.91), тышлыгында (13; 192.15), яңа (13; 10.95), тагын (13; 19.43), электрон (12; 62.8), кыйммәте (12; 122.25), редакторы (12; 109.51), ахырына (12; 77.81), рус (12; 40.54), тулыландырылган (11; 145.04), басылып (11; 66.24), үз (11; 4.91), эченә (11; 58.35), башында (11; 48.32), дөнья (11; 28.38), нинди (11; 21.16), берничә (11; 26.21), зурлыгы (11; 110.07), дәвами (10; 89.46), азагында (10; 69.35), икә (10; 7.43), эчендә (9; 23.93), аерым (9; 21.22), беренчесә (9; 61.76), баясә (9; 45.17), күп (9; 4.68), тарихы (9; 38.27), үзен (9; 30.2), тышлыгын (9; 143.19), исемәнә (9; 52.86), өченчә (9; 31.36), дәрәжәсен (8; 48.51), рәссамы (8; 66.54), үзенә (8; 23.84), чыгуы (8; 57.27), тышлыгына (7; 98.96), идея (7; 47.72), буеннан-буена (7; 78.26), илаһи (7; 49.64), тулы (7; 19.42), тышына (7; 80.96), максаты (7; 28.99), кулъязмасы (7; 74.51), тәржемәсен (7; 82.86), концепциясә (7; 61.05), дүртенчә (7; 35.75), ролән (6; 34.36), алгы (6; 40.17), берсен (6; 33.5), тәүгә (6; 28.69), битләрен (6; 57.76), баясен (6; 41.32), нигезенә (6; 48.04), урыны (6; 21.71), жаваплы (6; 23.14), нигезен (6; 43.45), авырлыгы (6; 44.28), барлык (6; 7.04), калган (6; 9.22), киңәйтелгән (5; 36.03), битләре (5; 44.65), тәржемәсә (5; 42.59), архив (5; 34.78), озакламый (5; 31.25), тышында (5; 56.4), баштагы (5; 41.35), кырыена (5; 45.43), россиягә (5; 34.23), эчтөлөгәндә (5; 61.1), рецензенты (5; 86) Барысын да карау

СУЛ КОНТЕКСТ: %^% (903; 1453.07), бу (305; 1270.4), , (224; 4.69), изгә (150; 1333.51), өлеге (119; 576.55), ул (69; 103.51), дигән (47; 150.31), исемлә (30; 171.36), яңа (27; 52.82)

Эксперименталь фонетика лабораториясе

Буялудан соң файдаланучы тәрәзәнең “**Мисаллар**” өлешендә ул “**китапның эчтәлеген**” сүзтезмәсе керткән җөмлөләргә тап булачак.

Әгәр “**китапның эчтәлеген**”нән соң килгән сүзләр ярдәмендә күпмәгънәлелек жуелмаса, шәхес “**китапның**” сүзе алдыннан килүче сүзләр белән дә танышырга мөмкин.



The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". There are navigation links for "ТАТ", "РУС", and "ENG". A sidebar on the left contains links like "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустан файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", and "Проект турында". Below the sidebar, there is contact information: "Элемтәгә керү: tatcorpus@gmail.com" and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". The main content area is titled "МИСАЛЛАР:" and lists several examples of the phrase "китапның эчтәлеген" in red text, each with a source citation and a "Текстны табу" link.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

МИСАЛЛАР:

- ▶ Аджтимушкай турындагы бу язманның яртысы диярлек Керчтә сатып алынган **китапның эчтәлеген** сөйләп чыгу булуга карастан, «сөрген» халкына бик сокландыргыч тәэсир ясады ул.
(чыганак: "Идел" журналы (web-сайт)) [Текстны табу](#)
- ▶ Элекке хәрбинең адвокаты исә **китапның эчтәлеген** рәсмиләр белән килештерергә кирәклеген кире кага.
(чыганак: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- ▶ **Китапның эчтәлеген** бик матур, сәнгатьле сөйли белгән ул.
(чыганак: "Яшел Үзән" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Ә Шөгәр төбәгендә гомер иткән Найлә Сиражетдин кызы Бикчуринаның язып калдырган истәлекләре, нәсел шәжәрәсе **китапның эчтәлеген** тарихи әһәмияткә ия мәгълүмат, документлар белән баетты.
(чыганак: "Самара региональ «Дуслык» һәм Самара шәһәр «Ак бәхет» ижади-ижтимагый оешмаларының уртак сайты" (web-сайт)) [Текстны табу](#)
- ▶ Китап исеме **китапның эчтәлеген** сөйләп бирергә тиеш түгел.
(чыганак: Рабит Батулла. Кәбир Бәкернең тууы)
- ▶ Шулай дип башлап китте дә ул кичә укыган **китапның эчтәлеген** сөйләргә тотынды.
(чыганак: Разил Вәлиев. Кунак булып килде, хужа булып китте...)
- ▶ Әйе, без синең хәтер күзәнәкләреңә йөз меңнәрчә **китапның эчтәлеген** язып куйдык.
(чыганак: Адлер Тимергалин. КОХАУ РОНГО-РОНГО)

Эксперименталь фонетика лабораториясе

2. Сул контекст аша. Моңың өчен шәхес “*китапның эчтәлеген*” сүзтезмәсенең сул контекстына мәрәжәгать итә. Файдаланучы бу очракта текст аланчыгына «*эчтәлеген*» сүзен яза һәм “Коллокацияләр (статистик) корпусында сүзформа эзләү...” дигән юл башындагы кнопкага тамга куя. Нәтижәдә яңа тәрәзә ачыла.

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area displays the search results for the word "эчтәлеген". The search input field and the "Табу!" button are highlighted with a red box. Below the search results, there are three radio button options, with the first one selected and highlighted by a red box:

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *лөргө**)

The footer of the page reads "Эксперименталь фонетика лабораториясе".

Тәрәзәдә “Сул контекст”тагы сүзләр арасында “*китапның*” сүзен табып аңа сул кнопка белән чиртә:

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". There are navigation tabs for "ТАТ", "РУС", and "ENG". On the left, there are several menu items: "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустаң файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтөгә керү:" (with email `tatcorpus@gmail.com`), and "Авторлар:" (listing Сайхуннов М.Р. and Ибраһимов Т.И.).

The main content area displays search results for the word "китапның". The "УҢ КОНТЕКСТ:" (Right Context) section lists various words with their frequency and frequency per million (e.g., "тәшкил (71; 568.22)", "һәм (65; 60.1)", "сөйләп (62; 609.69)", "дә (53; 78.42)", "яңарту (42; 473.06)", "сөйләү (38; 452.26)", "тагын (35; 143.58)", "ачып (33; 255.32)", "аңлау (21; 215.15)", "билгели (19; 171.77)", "аңлап (17; 129.27)", "ачу (16; 94.23)", "тулысынча (16; 101.47)", "яхшырту (14; 116.92)", "сөйләргә (14; 120.56)", "кыскача (14; 130.53)", "үзгәртү (14; 131.99)", "баету (14; 169.71)", "ача (13; 104.56)", "искә (12; 52.02)", "сөйли (12; 65.12)", "баета (11; 129.14)", "үз (10; 9.06)", "язып (10; 52.7)", "ачыклау (9; 65.94)", "камилләштерү (9; 73.21)", "дәрәс (9; 29.53)", "табу (9; 60.23)", "аңлату (9; 74.93)", "аңлатып (9; 67.58)", "үзгәртәргә (8; 67.1)", "ничек (8; 17.93)", "аңларга (8; 56.47)", "үзләштерү (7; 59.95)", "аңлый (7; 43.1)", "яхшы (7; 11.76)", "хасил (7; 58.28)", "шактый (6; 16.77)", "төгәл (6; 30.11)", "бәян (6; 38.33)", "ачарга (6; 40.23)", "тирәнәйтәргә (6; 91.36)", "сөйләгез (6; 66.54)", "тулырак (6; 46.85)", "әйтәп (6; 20.64)", "табабыз (5; 55.02)", "мәктәптә (5; 19.67)", "үзәнчәлекле (5; 23.57)", "ачуда (5; 45.24)", "тикшереп (5; 31.86)", "югалта (5; 40.28)", "кыска (5; 23.24)", "аңлауга (5; 59.92)", "ачар (5; 52.74)", "баетырга (5; 59.78)", "язу (5; 27.21)", "сайлау (5; 18.86)", "эшләү (5; 19.28)", "билгеләү (5; 29.75)", "бермә-бер (5; 41.76)", "тулыландырып (4; 39.82)", "бирә (4; 7.02)", "болай (4; 10.56)", "ачуга (4; 33.13)", "язарга (4; 20.58)", "истәлекләр (4; 30.09)", "саклаган (4; 30.5)", "яттан (4; 32.53)", "үзгәртә (4; 32.33)", "тулыландыру (4; 38.61)", "тирән (4; 15.33)", "саклап (4; 13.36)", "баетып (4; 42.29)", "халыкка (4; 13.77)", "тулы (4; 10.48)", "табарга (4; 19.72)", "сөйләтү (4; 60.72)", "аңлата (4; 19.67)", "сөйлиләр (4; 27.55)", "өйрәнә (4; 22.21)", "белү (4; 20.17)", "баетуга (4; 46.86)", "укуп (4; 13.67)", "ачыклай (4; 33.08)", "төкәдим (4; 5.24)", "; (4; 4.68)", "түбәндәгә (3; 14.01)", "алдагы (3; 10.51)", "телдән (3; 22.33)", "сөйләде (3; 8.55)", "үзгәртәп (3; 14.58)", "беләргә (3; 11.78)", "яңартуга (3; 27.54)", "авторның (3; 17.9)", "микән (3; 11.53)", "яхшырак (3; 14.17)", "салырга (3; 14.49)", "баеткан (3; 34.99)", "яңадан (3; 8.48) Барысын да карау

The "СУЛ КОНТЕКСТ:" (Left Context) section lists words with their frequency and frequency per million (e.g., "аның (114; 521.98)", "әсәрнең (91; 1154.86)", "төп (72; 466.93)", "һәм (59; 46.83)", "идея (59; 710.65)", "кыскача (45; 525.52)", "төкәтнең (41; 640.25)", "аларның (38; 173.86)", "китапның (26; 285.34)", "әсәрләренә (25; 312.73)", "әсәр (25; 215.59)", "хатның (22; 311.81)", "бирүнең (21; 235.2)", "биру (18; 82.78)", "тулы (14; 69.61)", "газетаның (14; 137.22)", "жырның (13; 146.04)", "хикәянең (13; 182.25)", "китапларның (12; 138.96)", "спектакльнең (11; 116.53)", "төкәт (11; 104.1)", "мәгънәви (9; 97.29)", "мәкаләнен (9; 112.41)", "оригиналның (8; 128.65)).

The word "китапның" in the left context is highlighted with a red box. At the bottom of the page, it says "Эксперименталь фонетика лабораториясе".

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит
 Корпуста эзләү
 Статистик мәгълүмат
 Корпус тасвирламасы
 Корпустаң файдалану
 Басылган мәкаләләр
 Рәхмәт сүзе
 Проект турында

Элемтәгә керү:
 tatcorpus@gmail.com

Авторлар:
 Сайхунов М.Р.
 Ибраһимов Т.И.

УҢ КОНТЕКСТ: , (146; 5.21), төшкил (71; 568.22), һәм (65; 60.1), сөйләп (62; 609.69), дә (53; 78.42), яңарту (42; 473.06), сөйләү (38; 452.26), тагын (35; 143.58), ачып (33; 255.32), аңлау (21; 215.15), билгели (19; 171.77), аңлап (17; 129.27), ачу (16; 94.23), тулысынча (16; 101.47), яхшырту (14; 116.92), сөйләргә (14; 120.56), кыскача (14; 130.53), үзгәртү (14; 131.99), баету (14; 169.71), ача (13; 104.56), искә (12; 52.02), сөйли (12; 65.12), баета (11; 129.14), үз (10; 9.06), язып (10; 52.7), ачыклау (9; 65.94), камилләштерү (9; 73.21), дәрәс (9; 29.53), табу (9; 60.23), аңлату (9; 74.93), аңлатып (9; 67.58), үзгәртәргә (8; 67.1), ничек (8; 17.93), аңларга (8; 56.47), үзләштерү (7; 59.95), аңлый (7; 43.1), яхшы (7; 11.76), хасил (7; 58.28), шактый (6; 16.77), төгәл (6; 30.11), бәян (6; 38.33), ачарга (6; 40.23), тирәнәйтәргә (6; 91.36), сөйләгез (6; 66.54), тулырак (6; 46.85), әйтәп (6; 20.64), табабыз (5; 55.02), мәктәптә (5; 19.67), үзәнчәлекләр (5; 23.57), ачуда (5; 45.24), тикшерәп (5; 31.86), югалта (5; 40.28), кыска (5; 23.24), аңлауга (5; 59.92), ачар (5; 52.74), баетырга (5; 59.78), язу (5; 27.21), сайлау (5; 18.86), эшләү (5; 19.28), билгеләү (5; 29.75), бермә-бер (5; 41.76), тулыландырып (4; 39.82), бирә (4; 7.02), болай (4; 10.56), ачуга (4; 33.13), язарга (4; 20.58), истәлекләр (4; 30.09), саклаган (4; 30.5), яттан (4; 32.53), үзгәртә (4; 32.33), тулыландыру (4; 38.61), тирән (4; 15.33), саклап (4; 13.36), баетып (4; 42.29), халыкка (4; 13.77), тулы (4; 10.48), табарга (4; 19.72), сөйләтү (4; 60.72), аңлата (4; 19.67), сөйлиләр (4; 27.55), өйрәнә (4; 22.21), белү (4; 20.17), баетуга (4; 46.86), укып (4; 13.67), ачыклай (4; 33.08), төкәдим (4; 5.24), ; (4; 4.68), түбәндәгә (3; 14.01), алдагы (3; 10.51), телдән (3; 22.33), сөйләде (3; 8.55), үзгәртәп (3; 14.58), беләргә (3; 11.78), яңартуга (3; 27.54), авторның (3; 17.9), микән (3; 11.53), яхшырак (3; 14.17), салырга (3; 14.49), баеткан (3; 34.99), яңадан (3; 8.48) Барысын да карау

СУЛ КОНТЕКСТ: , (179; 26.12), аның (114; 521.98), әсәрнең (91; 1154.86), төп (72; 466.93), һәм (59; 46.83), идея (59; 710.65), кыскача (45; 525.52), текстның (41; 640.25), аларның (38; 173.86), **китапның** (26; 285.34), әсәрләренң (25; 312.73), әсәр (25; 215.59), хатның (22; 311.81), бирүнең (21; 235.2), биру (18; 82.78), тулы (14; 69.61), газетаның (14; 137.22), жырның (13; 146.04), хикәянең (13; 182.25), китапларның (12; 138.96), спектакльнең (11; 116.53)

Эксперименталь фонетика лабораториясе

Тәрәзәнең “Мисаллар” исемле өлешендә шәхескә “*китапның эчтәлеген*” сүзтезмәсен эченә алган җөмлөләр исемлеге бирелә, шул чиратта “*китапның*” алдыннан килүче сүзләр белән танышу мөмкинлеге барлыкка килә.

The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content is under the heading "МИСАЛЛАР:" (Examples). There are seven examples, each starting with a play button icon and followed by a text snippet and a source citation. The phrase "китапның эчтәлеген" is highlighted in red in all examples.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

МИСАЛЛАР:

- ▶ Аджитмушкый турындагы бу язманның яртысы диярлек Керчтә сатып алынган **китапның эчтәлеген** сөйләп чыгу булуга карамастан, «сөрген» халкына бик сокландыргыч тәэсир ясады ул.
(чыганак: "Идел" журналы (web-сайт)) [Текстны табу](#)
- ▶ Элекке хәрбинең адвокаты исә **китапның эчтәлеген** рәсмиләр белән килештерергә кирәклеген кире кага.
(чыганак: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- ▶ **Китапның эчтәлеген** бик матур, сәнгатьле сөйли белгән ул.
(чыганак: "Яшел Үзән" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Ә Шөгәр төбәгендә гомер иткән Наилә Сиражетдин кызы Бикчуринаның язып калдырган истәлекләре, нәсел шәжәрәсе **китапның эчтәлеген** тарихи әһәмияткә ия мәгълүмат, документлар белән баетты.
(чыганак: "Самара региональ «Дуслык» һәм Самара шәһәр «Ак бәхет» иҗади-иҗтимагый оешмаларының уртак сайты" (web-сайт)) [Текстны табу](#)
- ▶ Китап исеме **китапның эчтәлеген** сөйләп бирергә тиеш түгел.
(чыганак: Рабит Батулла. Кәбир Бәкернең тууы)
- ▶ Шулай дип башлап китте дә ул кичә укыган **китапның эчтәлеген** сөйләргә тотынды.
(чыганак: Разил Вәлиев. Кунак булып килде, хужа булып китте...)
- ▶ Әйе, без синең хәтер күзәнәкләреңә йөз меңнәрчә **китапның эчтәлеген** язып куйдык.
(чыганак: Адлер Тимергалин. КОХАУ РОНГО-РОНГО)

Эксперименталь фонетика лабораториясе

3. **Катлаулы морфологик эзлөү системасы.** Өгәр дә сезгә статистик мәгълүмат кирәк булмаса, нәм сез бары тик шул сүзтезмәләр кәргән җөмлөләр исемлеген генә күрергә телисез икән, моның өчен иң җайлы юл - киләсе бүлектә тасвирланган [Катлаулы морфологик эзлөү системасын](#) куллану.

Катлаулы морфологик эзләү системасы:

Өлеге системаның асылы турында

2014 елда Корпусны морфологик тамгалау эше башкарылды. Грамматик тамгалар метателе нигезенә [Апертиум](#) халыкара проекты тарафыннан төрки телләр өчен эшлэнгән «тег»лар системасы алынды. Ул проект кысаларында дөньяның күп телләрендөгә текстларны автоматик рәвештә үзара тәржемә итү системасы эшләнә. Язма корпусны тамгалау өчен нәкъ менә Apertium'ны сайлау түбәндөгә сәбәпләр белән аңлатыла:

- морфологик аннотациянең югары дәрәжәдә булуы;
- ул проектның тулысынча ачыклығы, ягъни бөтен код һәм мөмкинлекләренең теләсә кем тарафыннан бушлай кулланыла алуы.

2015-2016 елларда безнең тарафтан эшлэнгән **Катлаулы морфологик эзләү системасы** Корпуста сүзформа, лемма, морфологик (грамматик) теглар, сүзнең башы, уртасы, азагы, алар арасындагы ераклык кебек билгеләрнең төрле комбинацияләре нигезендә мисаллар эзләргә мөмкинлек бирә. Эзләнәсе юлның озынлыгы биш сүзгә җитә, һәм алар арасындагы дүрт ераклык күрсәткече дә каралган.

Катлаулы морфологик эзләү формасын табар өчен сул яктагы “Корпуста эзләү” ссылкасына чиртегез һәм ачылган тәрәзедә “Катлаулы морфологик эзләү” юлын сайлап алыгыз:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

баш биткә

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<input type="text"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>			<input type="button" value="Найти!"/>

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

“Кунакка бардык” тезмәсен эзләп карыйк:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

баш биткә

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр






Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

Сүз 1:	<input type="text" value="кунакка"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text" value="бардык"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>			<input type="button" value="Найти!"/>

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

Эзләү нәтижеләре битенәң иң өске өлешендә эзләү башкарылган параметрлар китерелә. Аның астында Корпуста табылган очракларның төгәл саны бирелә.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

ЭЗЛӘНӘСЕ ЮЛ: {кунакка} 1-1 {бардык} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 20

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

- ▶ Икенче көн дә күтәренке күңел белән узды: хатынның туганнарына **кунакка бардык**, яисә үзләре безгә килделәр.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Менә без дә Чаллыга **кунакка бардык**, шәһәр карап, аның матурлыгына сокланып йөрдек.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Шактый еллар узгач, аның илле яшьлек юбилеена **кунакка бардык**.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ – Кунаклар килде, **кунакка бардык**, – дип җавап кайтаралар.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Иртүк мөезин хатынына **кунакка бардык**.
(чыганақ: "Идел" журналы (web-сайт)) [Текстны табу](#)
- ▶ Без әни белән Нәфисә апаларга **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Бервакыт Украинада яшәүче әтиемнең туганнарына **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Мәсәлән, узган атнада җәйге лагерыбызда тәрбияләнүче балаларыбыз белән аларга **кунакка бардык**.
(чыганақ: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- ▶ Без, делегация белән, Яр Чаллы каласына Советлар Союзы Герое Илдар абый Маннанов янына **кунакка бардык**.
(чыганақ: "Әдилләр. Әхмәт Дусайлының шәхси порталы" (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Эзләү башында күрсәтелгән параметрлар табылган жөмлөлөрдә кызыл төстө бирелә. Тычканның курсорын шуларның һәркайсына китергөч (чиртмиçә), махсус тәрәзәдә әлеге сүзнең леммасы (Apertium системасы интерпретациясендә) һәм аңа хас морфологик тамгалар жыелмасы бирелә.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Файдалы мәгълүмат

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

ЭЗЛӘНӘСӘ ЮЛ: {кунакка} 1-1 {бардык} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 20

- ▶ Икенче көн дә күтәрәнке күңел белән узды: хатынның туганнарына **кунакка бардык**, яисә үзләре безгә килделәр.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Менә без дә Чаллыга **кунакка бардык**, шәһәр карап, аның матурлығына сокланып йөрдек.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Шактый еллар узгач, аның илле яшьлек юбилеена **кунакка бардык**.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

(кунак) <dat>.<n>.<sg>
- ▶ – Кунаклар килде, **кунакка бардык**, – дип җавап кайтаралар.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Иртүк мөезин хатынына **кунакка бардык**.
(чыганақ: "Идел" журналы (web-сайт)) [Текстны табу](#)
- ▶ Без әни белән Нәфисә апаларга **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Бервакыт Украинада яшәүче әтиемнең туганнарына **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Мәсәлән, узган атнада җәйге лагерыбызда тәрбияләнүче балаларыбыз белән аларга **кунакка бардык**.
(чыганақ: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- ▶ Без, делегация белән, Яр Чаллы каласына Советлар Союзы Герое Илдар абый Маннанов янына **кунакка бардык**.
(чыганақ: "Әдилләр. Әхмәт Дусайлының шәхси порталы" (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Тычкан курсорын “бардык” сүзенә китергәч түбәндәге мәгълүмат күрсәтелә:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

ЭЗЛӘНӘСӘ ЮЛ: {кунакка} 1-1 {бардык} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 20

- ▶ Икенче көн дә күтәрәнке күңел белән узды: хатынның туганнарына **кунакка бардык**, яисә үзләре безгә килделәр.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Менә без дә Чаллыга **кунакка бардык**, шәһәр карап, аның матурлыгына сокланып йөрдек.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Шактый еллар узгач, аның илле яшьлек юбилеена **кунакка бардык**.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ – Кунаклар килде, **кунакка бардык**, – дип җавап кайтаралар.
(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Иргүк мөезин хатынына **кунакка бардык**.
(чыганақ: "Идел" журналы (web-сайт)) [Текстны табу](#)
- ▶ Без әни белән Нәфисә апаларга **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Бервакыт Украинада яшәүче әтиемнең туганнарына **кунакка бардык**.
(чыганақ: "Яңарыш" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Мәсәлән, узган атнада җәйге лагерьыбызда тәрбияләнүче балаларыбыз белән аларга **кунакка бардык**.
(чыганақ: "Азатлык Радиосы" (web-сайт)) [Текстны табу](#)
- ▶ Без, делегация белән, Яр Чаллы каласына Советлар Союзы Герое Илдар абый Маннанов янына **кунакка бардык**.
(чыганақ: "Әдиләр. Әхмәт Дусайлының шәхси порталы" (web-сайт)) [Текстны табу](#)

(бар) <ifi>, <iv>, <p1>, <pl>, <v>

Эксперименталь фонетика лабораториясе

Истә тотыгыз: Корпусны морфологик тамгалау автоматик рәвештә башкарылу сәбәпле, хаталарның очравы ихтимал!

Корпуста сүзформаларны үз эченә алган комбинацияләр нигезендә эзләү

Корпуста теге яки бу сүзформаларның тезмәсен табу өчен “Сүз 1”, “Сүз 2”, “Сүз 3”, “Сүз 4”, “Сүз 5” дигән текст аланчыкларына кирәкле сүзләрне берәм-берәм языгыз. Шулай итеп, эзләнәсе тезмәнең озынлыгы биш сүзгә кадәр җитәргә мөмкин.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

баш биткә

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<input type="text" value="яңа"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text" value="ел"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text" value="бәйрәмендә"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>		<input type="button" value="Найти!"/>	

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

ЭЗЛӘНӘСЕ ЮЛ: {яңа} 1-1 {ел} 1-1 {бәйрәмендә} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 55

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

▶ Алар **Яңа ел бәйрәмендә** балаларны һәм, гомумән, шундагы һәркемне чын Кыш бабай һәм Кар кызы белән очрашу оештыралар.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ Инде китү ягына борылгач та, Наилә апа артымнан килеп: "Кызым, тагын кил. **Яңа ел бәйрәмендә** бик күңелле була", – дип кочып алды.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ Бу – **Яңа ел бәйрәмендә** 600 млн литр исерткеч елга булып агачак дигән сүз.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ **Яңа ел бәйрәмендә** һәм Сабантуйда.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ **Яңа ел бәйрәмендә** уңышлыгына чыгыш ясадык.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ **Яңа ел бәйрәмендә** бик күңелле була", – дип кочып алды.

(чыганақ: "Безнең гәҗит" газетасы (web-сайт)) [Текстны табу](#)

▶ **Яңа ел бәйрәмендә,**

(чыганақ: "Туган як" газетасы (web-сайт)) [Текстны табу](#)

▶ – Эдуард Анатольевич, халыкны **Яңа ел бәйрәмендә** чыршылар белән тәэмин итүгә әзерлек мәсьәләсе ничек тора?

(чыганақ: "Зәй офыклары" газетасы (web-сайт)) [Текстны табу](#)

▶ Республика **Яңа ел бәйрәмендә** катнашучыларны тәбрикләргә Татарстан Президенты Рөстәм Миңнеханов килде.

(чыганақ: Дәүләт хакимияте һәм җирле үзидарә органнарының бердәм "Рәсми Татарстан" порталы (web-сайт)) [Текстны табу](#)

Өгәр дә сезгә янәшә торган гына түгел, ә бер-берсеннән башка сүзләр белән бүленгән сүзләр дә эзләргә кирәк булса, моның өчен язылган сүзнең уң ягындагы “**Ара**” дигән аланчыкка кирәкле санны кертә аласыз. Мәсәлән, “**1-1**” яки “**1**” язсагыз, сүзләр янәшә килергә тиеш булалар, “**2-2**” яки “**2**” исә бу сүзләр арасында бер сүз булырга тиеш дигәнне аңлата һәм алга таба шул тәртиптә. “**1-5**” дип күрсәтелгән очракта, эзләнәсе сүзләр янәшә дә килергә мөмкин, алар арасында бер, ике, өч, дүрт сүз дә булырга мөмкин. Сүзләр арасындагы ераклык билгесе техник яктан чикләнмәгән, ул бары тик жөмлө озынлыгы белән генә билгеләнә.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	бик		Ара 1:	2-4
Сүз 2:	алдылар		Ара 2:	1-1
Сүз 3:			Ара 3:	1-1
Сүз 4:			Ара 4:	1-1
Сүз 5:				Найти!

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

ЭЗЛӘНӘСЕ ЮЛ: {бик} 2-4 {алдылар} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 179

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустаң файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

▶ Японнар тупны һәм уен кырының үзәген **бик** тиз яулап **алдылар**.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Безнең спортчыларыбыз монда яхшы чыгыш ясый, ләкин алар результатларын яхшыртуны күздә тоталар һәм шуның өчен дә монда **бик** шәп тәҗрибә **алдылар**.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Аккредитация үзәгендә безне **бик** ачык каршы **алдылар**, документларны тикшерделәр, банк картасы белән бейджлар тапшырдылар.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Корылманың персоналы күзәтүчеләр төркемен **бик** мөлаем каршы **алдылар** һәм үз объектларындагы функцияләре турында бик теләп сөйләде.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ "Безне монда **бик** әйбәт каршы **алдылар**, хезмәт күрсәтү бик югары дәрәҗәдә оештырылган.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ "Кызларга рәхмәт, **бик** яхшы каршы **алдылар**.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Ут Эстафетасының фирма поезды Курск вокзалына килде, монда Бөтендөнья студентлар Уеннарының төп символын мәскәүлеләр фанфаралар тавышы астында **бик** жылы каршы **алдылар**.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Йөзләргә төмәнлеләр Эстафета колоннасын бар хәрәкәт маршрутында **бик** жылы каршы **алдылар**.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Атташе булу өчен чит телләргә әйбәт белү генә аз, шуңа да Универсиада өчен аларны **бик** жентекләп сайлап **алдылар**.

Эксперименталь фонетика лабораториясе

Корпуста леммаларны үз эченө алган комбинацияләр нигезендө эзләү

Сез эзләүне билгеле бер сүзформа нигезендө түгел, ә аерым лемманың бөтен формалары нигезендө башкарырга телисез икән, моның өчен лемманы жәяләр эчендө язарга кирәк, мәсәлән, “(китап)”. Система бу лемманы базада таба алмаса, ул аны сүзформа буларак эзләячөк.

Бу юлны эзләп карыйк: “{теге} {(китап)}”:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

Сүз 1:	<input type="text" value="tere"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text" value="(китап)"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>		<input type="button" value="Найти!"/>	

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

ЭЗЛӘНӘСЕ ЮЛ: {теге} 1-1 {китап} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 22

- ▶ - Булат, **теге китап** сиңдәме әле? - диде Харис.
(чыганақ: Заһит Мурсиев. Ачылмаган тәрәзә)
- ▶ Балачагым күз алдыма килде, "Путивльдән Карпатка кадәр" дигән **теге китап та** искә төште.
(чыганақ: "Роберт Миңнуллин (официаль сайт)" (web-сайт)) [Текстны табу](#)
- ▶ Әмма **теге китап** күңелдән китми генә.
(чыганақ: "Роберт Миңнуллин (официаль сайт)" (web-сайт)) [Текстны табу](#)
- ▶ Тагын туп-туры **теге китап** торган бүлеккә ашыгасың.
(чыганақ: "Роберт Миңнуллин (официаль сайт)" (web-сайт)) [Текстны табу](#)
- ▶ Кави, син минем **теге китапны** укыдыңмы?
(чыганақ: Рабит Батулла. Тузга язылган хәлләр (Тарихи шәхесләрбез, замандашларыбыз турында мәзәкләр, хикәвләр, фа Жигале хәбәрләр))
- ▶ Кави, син минем **теге китапны** укыдыңмы?
(чыганақ: Рабит Батулла. Тузга язылган хәлләр (Тарихи шәхесләрбез, замандашларыбыз турында мәзәкләр, хикәвләр, фа Жигале хәбәрләр))
- ▶ Хәер, теге төрек тарихын, **теге китапны** биш-алты кат укып чыктым.
(чыганақ: Гаяз Исхакий. Тормышмы бу?)
- ▶ - Каенанам хатынга, кияү **теге китабын** биреп торсын әле, мин дә укып чыгар идем, дигән.
(чыганақ: "Мәдәни җомга" газетасы (web-сайт)) [Текстны табу](#)
- ▶ Ярты сәгать тә үтми, томшыгына кыстырып **теге китап** битен алып та килә.
(чыганақ: "Татарча текстлар, программалар" (web-сайт)) [Текстны табу](#)
- ▶ **Теге китап**, дәфтәрләр энә шул атларның аяк астында тапталып ятканнар.
(чыганақ: "Әдипләр. Әхмәт Дусайлының шәхси порталы" (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Башка мисал рәвешендә “{(авыл)} 1-3 {(кил)}” очрагын карыйк. Монда эзләү ике лемма нигезендә башкарыла, аларның арасы ике сүзгә кадәр житәргә мөмкин.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ





ТАТ РУС ENG

баш биткә

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1: (авыл)		Ара 1: 1-3
Сүз 2: (кил)		Ара 2: 1-1
Сүз 3:		Ара 3: 1-1
Сүз 4:		Ара 4: 1-1
Сүз 5:		Найти!

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

ЭЗЛӘНӘСЕ ЮЛ: {авыл} 1-3 {кил} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 4302

▶ – Мәсәдә **авылына** язучылар **килгән**, барабызмы?

(чыганак: Бари ИСЛАМ, БИК ГАДИ ДӘ, СЕРЛЕ ДӘ СИН, ТОРМЫШ)

▶ Чөнки якын тирә **авылларга килеп** утырган Гәрәйләр, Меңнәр, Кара бәк нәселләренең шәжәрәләрендәгә кешеләр исемнәре бу авылларда әле дә байтак.

(чыганак: Илдус Хужин, Топонимнар)

▶ Урта Баяк турында 17 гасыр документларында мәгълүматлар юк, ахрысы **авыл** соңрак барлыкка **килгән**.

(чыганак: Илдус Хужин, Топонимнар)

▶ Бу очракта **авылының** исеме **килеп** чыгышы башка күп кенә авылларныкына туры килә, чөнки Кантуган дигән шәхес тарихта бар.

(чыганак: Илдус Хужин, Топонимнар)

▶ Әлбәттә, коңгырат кабиләләреннән тыш бу **авылда** борынгыдан **килгән** төрки ырулар яшәгәндер һәм дә 400-500 еллар элек көчле гәрәйләр һәм Кара бәк кабиләләре анда килеп утырып яңа исем биргәннәр булырга тиеш.

(чыганак: Илдус Хужин, Топонимнар)

▶ Әйтергә кирәк, урыслар **авылга** соңрак **килеп** урнашкан булырга тиешләр, чөнки алар татар авылы белән аның зираты арасына кереп утырганнар.

(чыганак: Илдус Хужин, Топонимнар)

▶ Шулай итеп Уралда кыпчакларның һәм Кара бәк нәселенең бик күп **авыллары** дөньяга **килә**.

(чыганак: Илдус Хужин, Топонимнар)

▶ Ә документлар күрсәтүенчә, 1647 елда Өфе елгасы буендагы Өфе Шигер һәм Өртә Шигер **авыллары** барлыкка **килгән**.

(чыганак: Илдус Хужин, Топонимнар)

Корпуста морфологик (грамматик) тегларны үз эченә алган комбинацияләр нигезендә эзләү

Язма Корпуста кулланылган теглар исемлегә белән түбәндәге сылгалар буенча күчеп таныша аласыз:

1. http://corpus.tatar/index.php?openinframe=manual/tags_uniq.pdf
2. <https://sourceforge.net/p/apertium/svn/61954/tree//branches/turkic/lexc.rkt>

Эзләү бер һәм берничә тег нигезендә башкарылырга мөмкин. Бары тик “һәм” (“яки” түгел) төрендәге комбинацияләр генә кулланыла ала, бер-берсенә каршы торган парадигмаларны (мәсәлән, “<dat><acc>” *юнәлеш килеше, төшем килеше*) берьюлы эзләү мөмкин түгел.

Өйдәгез, “{<adj>} 1-2 {<n><dat>} 1-3 {<v><past>}” жыелмасын эзләп карыйк. Монда беренче сүз *сыйфат* (<adj>) булырга тиеш, аңардан соң *бер сүзгә* кадәр арада *исем сүз төркеменә* (<n>) караган берәмлек килә, ул *юнәлеш килешендә* булырга тиеш (<dat>), шуннан соң *ике сүзгә* кадәрге арада *фигыль* (<v>) килә, ул «-ган / -гән / -кан / -кән» *үткән заман формасында* (<past>).

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

[баш биткә](#)

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

Сүз 1:	<input type="text" value="<adj>"/>		Ара 1:	<input type="text" value="1-2"/>
Сүз 2:	<input type="text" value="<n><dat>"/>		Ара 2:	<input type="text" value="1-3"/>
Сүз 3:	<input type="text" value="<v><past>"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>		<input type="button" value="Найти!"/>	

Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)

Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)

Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)

Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

ЭЗЛӘНӘСЕ ЮЛ: {<adj>} 1-2 {<n>,<dat>} 1-3 {<v>,<past>} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 10594

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

▶ Авылларыбызның килеп чыгуы вакыты турында фикер йөрткәндә, шуларны онытмау мәгъкуль: башкорт галимнәренең язмалары буенча, бөтен татар авыллары да 1650 – 1750 елларда һәм **соңрак барлыкка килгәннәр**.

(чыганак: Илдус Хужин. Топонимнар)

▶ Урта Бәяк турында 17 гасыр документларында мәгълүматлар юк, ахрысы авыл **соңрак барлыкка килгән**.

(чыганак: Илдус Хужин. Топонимнар)

▶ **Соңрак бу урынга** башкалар да **килгән**.

(чыганак: Илдус Хужин. Топонимнар)

▶ Бу авыл башкортлары (ягъни жирбиләүчеләре, вотчинники) **соңрак** Әртә заводына жирләрен **сатканнар** һәм үзләре шул ук жирләрдә яшиләр дип язылган бер документта.

(чыганак: Илдус Хужин. Топонимнар)

▶ Авыл халкының байтагы күмер яндырып **якындагы заводларга** ташып **сатканнар**.

(чыганак: Илдус Хужин. Топонимнар)

▶ **Ак шәлләре жиргә** шуып **төшкән...**

(чыганак: Разил Вәлиев. Шигырьләр, Жырлар)

▶ 21 яшьлек шагыйрь 1907 елда үз арамыздан чыккан маһир рәссам хакында хыялланганда, андый рәссам инде хәзерге Татарстанның Буа районы (әлеккеге Тәтеш өязе) Күл Черкене авылында 1897 елның 22 февралендә Мәхжүбә абыстай һәм Идрис хәзрәт Урманчелар гаиләсендә **якты дөньяга** килгән **булган**.

(чыганак: "Татарстан китап нәширияты" (web-сайт)) [Текстны табу](#)

▶ Соңрак рус тарихчылары, мәгънәне берәз тарайтып, барлык төркиләргә татар дип атаганнар, **төрки** сүзе **урынына** татар сүзен **кулланганнар**.

(чыганак: "Татарстан китап нәширияты" (web-сайт)) [Текстны табу](#)

Түбөндөгө мисалда янәшә торган ике сүз карала. Аларның беренчесе *икенче зат берлек сан тартым (<px2sg>)* белән төрлөнгән, ә икенчесе *төшем килешендә (<acc>)* булырга тиеш.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

Сүз 1:	<input type="text" value="<px2sg>"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text" value="<acc>"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>		<input type="button" value="Найти!"/>	

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 нәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatorpus@gmail.com

Эксперименталь фонетика лабораториясе

TAT РУС ENG

ЭЗЛӘНӘСЕ ЮЛ: {<px2sg>} 1-1 {<acc>} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 4954

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мөкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Авторлар:

Сайхунов М.Р.

Ибраһимов Т.И.

tatcorpus@gmail.com

▶ Шуңа күрә, күптән инде икенче халыкка әйләнгән үз **халкыңның эзләрен** эзләү файдасыз.

(чыганақ: Илдус Хужин. Топонимнар)

▶ Менә шул үзсүзлегең, кеше сүзен, **анаң сүзен** тыңламавың харап итә яздымы үзеңне?

(чыганақ: Фәрит Яхин. СЕРЛЕ КАМКА)

▶ Була шундый чаклар, син якын **дустыңның ялганын** тыңдыйсың.

(чыганақ: "Татарстан китап нәширияты" (web-сайт)) [Текстны табу](#)

▶ Такта өстенә **жүм саласың да ятьмәне** бакчага чыгарып куясың.

(чыганақ: "Татарстан китап нәширияты" (web-сайт)) [Текстны табу](#)

▶ Тик, зинһар өчен, шаярышмагыз анда, рәссам **абыйларыгызының кушканын** төгәл үтәгез!

(чыганақ: "Татарстан китап нәширияты" (web-сайт)) [Текстны табу](#)

▶ - Их сез, яшел чебешләр, яшь **чагыгызының кадерең** белеп калыгыз.

(чыганақ: Заһит Мурсиев. Мин сезнең хатыңыгыз)

▶ – И улым, болары гына аның бик вак мәсьәлә, артык **исең китмәсен**, менә үз көнеңне үзең күрә башлагач, дөнья әле уңны–суңны яңаклар...

(чыганақ: РАВИЛ ӘМИРХАН. ӘМИРХАННАР (тәфсилле шәһәрә). КАЗАН – 2005)

▶ Дүшәмбе көн безнең уртақ кичәдә өзәлгән **аягыңның тарихын** сөйләрсен, Зәңгәр чишмә буйлары мондый хәлләргә иштергә бик сусаганнар... ярыймы? — дип, миннән кат-кат вәгдәләр алгач, шаулашып чыгып киттеләр.

(чыганақ: ГАЛИМҖАН ИБРАҺИМОВ. КЫЗЫЛ ЧӘЧӘКЛӘР)

▶ Ә инде семинар актив аралашуны, балалар белән аралашу өчен үз **программаңны формалаштыруны**, яңа методлар, технологияләр табуны күздә тоткың, бу мөгаен, педагогика белеменең иң зур нәтижәседер.

(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Корпуста сүзнең башы, уртасы һәм/яки азагы күрсәткече кулланылган комбинацияләр нигезендә эзләү

Билгеле бер хәрефләр тезмәсенә башланган сүзләрне табу өчен “**ки***” кебек юл кулланыла. Нәтижәдә “**китап, китапны, киштәгә, кит...**” кебек сүзләр табылачак.

Сүзнең уртасы буенча эзләү өчен «***әме***» дип яза аласыз: **керәмен, әмер, үсәме...**

Сүз азагы нигезендә эзләү “***рны**” юлы ярдәмендә башкарыла. Җаваплар исемлегендә “**дусларны, тарны, барны, кулларны...**” кебек сүзләр кергән җөмлөләр булачак.

Өйдәгез, “**{ил*} 1-1 {белән}**” җыелмасын эзләп карыйк, ягъни беренче сүз “**ил**”гә башланырга, ә икенче сүз исә “**белән**”гә туры килергә тиеш була.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит

Корпуста эзләү

Татарча сөйләгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустан файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзне языгыз:

Сүз 1:	<input type="text" value="ил*"/>		Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text" value="белән"/>		Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>		Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>		Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>		<input type="button" value="Найти"/>	

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрефтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatorpus@gmail.com

Эксперименталь фонетика лабораториясе

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мөкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

ЭЗЛӘНӘСӘ ЮЛ: {ил} 1-1 {белән} 1-1 {} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 2695

▶ Элеге масштаблы проектны тормышка ашыру Татарстан башкаласын үстерүгә, аның Россия төбәкләре белән элемтәләрен ныгытуга зур этәргеч булды, халыкта патриотизм һәм **ил белән** хорурлык хисе артты.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Яңа ел театральләштерелгән тамашасында барс балаларының ничек итеп Универсиада Утын эзләп төрле континентлар буенча сәяхәт итүләре, чит **илләр белән** танышулары, төрле мажараларга тарулары, төрле һөнәрләргә өйрәнүләре һәм ниһаять, Универсиада Утын кулга төшерүләре тасвирланды.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Универсиадада Россия медальләр өчен Кытай, АКШ кебек **илләр белән** көч сынашачак.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Матди стимуллардан тыш, волонтерлар өчен бик киң күңел ачу программасы да әзерләнгән: Универсиада **Илчеләре белән** иртәнге аш һәм очрашулар, һәр кичтә МЕГАда кичәләр һәм Уеннар тәмамланганнан соң бик якты зур йомгаклау бәйрәме.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ Бу максат, әлбәттә, башка **илләр белән** очрашуда да кабатлана.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ - Универсиада **илчеләре белән** килешүләре ничек бара?
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ 2013 елгы Универсиаданың кайда үтәчәге турындагы карарны берничә ел элек бик дулкынланып, тын да алмый **илебез белән** көткән идек.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)

▶ "Казан 2013" волонтерлары өчен Универсиада **Илчеләре** белән күзгә-күз сөйләшү эш сменасына килү кебек үк гади нәрсә: нәкъ менә объектлардагы буш вакытларында егетләргә һәм кызларны теге яки бу спорт

Эксперименталь фонетика лабораториясе

Сүзнең башы һәм азагы нигезендә мисаллар табу өчен **“ал*лар”** дип язарга кирәк. **“Алмалар, алмагачлар, алдылар, алалар...”** кебек сүзләр кайтарылачак.

Сүзнең башы, уртасы һәм азагы буенча эзләү шәкеле - **«к*әме*н»**: **керәмен, каләмен, күләменнән, кияүдәмен...**

Йолдыз **«*»** билгесе теләсә ничә (хәтта нуль дә) символны белдереп килсә, сорау билгесе **«?»** теләсә кайсы бер символны алыштырып килә. Мәсәлән, **«т?з*»** язмасы **«тиз, тозны, түзде, тазарды...»** кебек очрақларга туры киләчәк, ләкин **«тигез, тугызны, төрәзә...»** сүзләренә туры килмәячәк.

Саналып киткән барлык эзләү параметрларын (сүзформа, лемма, грамматик теглар, сүз башы, сүз уртасы һәм сүз азагы) төрле юл аланчыкларында кулланып, теләсә нинди комбинацияләр төзәргә мөмкин. Мәсәлән, “{<prn>} 1-1 {(кеше)} 1-3 {ал*}” жыелмасы. Бу очракта беренче сүз *алмашлык* (<prn>), аның белән янәшә “кеше” *леммасы* (ягъни кеше сүзенең теләсә кайсы формасы: **кешеләр, кешенең, кешеләребезгә...**) килә. Шуннан соң ике сүзгә кадәр ераклыкта “ал” хәрәфләр тизмәсенә башланган сүз булырга тиеш:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

баш биткә

Баш бит
Корпуста эзләү
 Татарча сөйләгеч
Кулланма
 Статистик мәгълүмат
 Корпус тасвирламасы
 Корпустаң файдалану
 Басылган мәкаләләр
 Рәхмәт сүзе
 Проект турында
 Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<prn>	⌨	Ара 1:	1-1
Сүз 2:	(кеше)	⌨	Ара 2:	1-3
Сүз 3:	ал*	⌨	Ара 3:	1-1
Сүз 4:		⌨	Ара 4:	1-1
Сүз 5:		⌨	Найти!	

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)
- Катлаулы морфологик эзләү (Кулланма)

Эксперименталь фонетика лабораториясе

TAT РУС ENG

Баш бит
 Корпуста эзләү
 Татарча сөйләгеч
Кулланма
 Статистик мәгълүмат
 Корпус тасвирламасы
 Корпустан файдалану
 Басылган мәкаләләр
 Рәхмәт сүзе
 Проект турында

Файдалы мәгълүмат

Авторлар:
 Сайхунов М.Р.
 Ибраһимов Т.И.
tatcorpus@gmail.com

ЭЗЛӘНӘСЕ ЮЛ: {<prt>} 1-1 {кеше} 1-3 {ал} 1-1 {} 1-1 {}

ТАБЫЛУ САНЫ: 514

- ▶ Мин хаж и булгач инде, үзең беләсең **ул кешеләрнең** догаларын **алырга** кирәклегә тугрысында.
(чыганақ: РАВИЛ ӘМИРХАН. ӘМИРХАННАР (тәфсилле шәһәрә). КАЗАН – 2005)
- ▶ Хәзер инде **мин кешеләрне, аларның** үз-үзләрен тотышын өйрәндем.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)
- ▶ Үз күзләрең белән күргәннәреңне **шулкадәр кешеләргә** тапшыра **алуыңны** тою - бик тә рәхәт хис икән ул.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)
- ▶ Бүген бөтен Россия буенча факелчылар тарафыннан күтәрәп барылган һәм 30 зур шәһәрне узган Уеннарның төп символы Универсиада авылында урнашкан Универсиада музееда тора һәм **аны кешеләр** якыннан күрә **алалар**.
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)
- ▶ **Мин кешеләргә** ярдәм итә **алам** икән, нигә моннан файдаланмаса?
(чыганақ: "2013 Универсиадасы блогы" (web-сайт)) [Текстны табу](#)
- ▶ **Андый кеше Аллаһы Тәгалә** тарафыннан бирелә, лидер мохит эчендә тәрбияләнә ала.
(чыганақ: "Үзбездә" яңа буын хәрәкәте (web-сайт)) [Текстны табу](#)
- ▶ – **Нинди кеше** булсын **алар**?
(чыганақ: Жәүдәт Юныс. Сугыштан соң)
- ▶ Дәрәс, ул **болай да кешеләр** белән артык **алыш**-би-реш итми, аралашмый, аңа йомышы төшәп, йә тәрәзә рамнары, йә балта сабы, йә көянтә-мазар ясатырга килүчеләр булса да, үзе беркемгә йөрми, айга бер-ике кибет тирәсен урап кайта да, минем сезгә катнашым юк, дигәндәй, тагын атавына кереп бикләнә.
(чыганақ: Разил Вәлиев. Мирас)

Корпуста эзләнәсе һәр сүзгә берничә параметр күрсәтү

Эзләү системасында һәр сүз сүзформа, лемма, теглар, сүз башы, уртасы, азагы кебек төрле параметрлар белән тасвирланырга мөмкин.

Мисал буларак түбәндәге очракны карап үтик: безгә “алма” (*фигыль*) сүзформасын табарга кирәк. Ләкин эзләү нәтижәләре арасында “алма” (*исем*) кәргән жөмлөләр да булачак. Бары тик беренче очракны гына эзләү өчен безгә “алма<v>” дип язарга кирәк. Монда “<v>” теги *фигыль*не белдерә.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

баш биткә

ТАТ РУС ENG

Баш бит
Корпуста эзләү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<input type="text" value="алма<v>"/>			Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text"/>			Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>			Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>			Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>			<input type="button" value="Табу!"/>	

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны**, **авылларга**, **килмөдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап**, **авыл**, **кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап***, **авыл***, ***гез**, ***лөргө**)
- Катлаулы морфологик эзләү (Кулланма, Теглар исемлеге)

Эксперименталь фонетика лабораториясе

Өлеге максатка ирешүнең икенче юлы да бар. Икенче аргумент буларак лемманы күрсәтергә мөмкин: “алма(ал)”. Ягъни система, леммасы “ал” булган “алма” сүзформаларын гына исәпкә алачак, чөнки исем сүз төркеменә кергән “алма” сүзенең леммасы - “алма”.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит

Корпуста эзләү

Татарча сөйлөгеч

Кулланма

Статистик мәгълүмат

Корпус тасвирламасы

Корпустаң файдалану

Басылган мәкаләләр

Рәхмәт сүзе

Проект турында

Файдалы мәгълүмат

Татар теленең язма корпусында эзләү

Эзләү төрөн сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<input type="text" value="алма(ал)"/>			Ара 1:	1-1
Сүз 2:	<input type="text"/>			Ара 2:	1-1
Сүз 3:	<input type="text"/>			Ара 3:	1-1
Сүз 4:	<input type="text"/>			Ара 4:	1-1
Сүз 5:	<input type="text"/>			<input type="button" value="Табу!"/>	

Авторлар:

Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

- Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны**, **авылларга**, **килмәдегез**)
- Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап**, **авыл**, **кил**)
- Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап***, **авыл***, ***гез**, ***ләргә**)
- Катлаулы морфологик эзләү (Кулланма, Төглар исемлеге)

Эксперименталь фонетика лабораториясе

Эзләү параметрларын график режимда кертү

Төрле эзләү параметрларын кертүне жиңеләйтү өчен махсус инструмент бар. Аны ачу өчен текст аланчыкларының уң ягында урнашкан кызыл түгәрәк төймәгә чиртергә кирәк.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

баш биткә

Баш бит
Корпуста эзләү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatorpus@gmail.com

Татар теленең язма корпусында эзләү

Эзләү төрен сайлап, кирәкле сүзгә языгыз:

Сүз 1:	<input type="text"/>			Ара 1:	<input type="text" value="1-1"/>
Сүз 2:	<input type="text"/>			Ара 2:	<input type="text" value="1-1"/>
Сүз 3:	<input type="text"/>			Ара 3:	<input type="text" value="1-1"/>
Сүз 4:	<input type="text"/>			Ара 4:	<input type="text" value="1-1"/>
Сүз 5:	<input type="text"/>				<input type="button" value="Табу!"/>

Коллокацияләр (статистик) корпусында сүзформа эзләү (мәсәлән, **китапны, авылларга, килмәдегез**)

Морфологик тамгаланган корпуста лемма эзләү (мәсәлән, **китап, авыл, кил**)

Шаблон буенча эзләү (3 һәм артыграк хәрәфтән торган сүз башы яки азагы күрсәтелә, мәсәлән, **китап*, авыл*, *гез, *ләргә**)

Катлаулы морфологик эзләү (Кулланма, Теглар исемлеге)

Эксперименталь фонетика лабораториясе










Рәсемдә күрсәтелгән тәрәзә ачылачак:

ТАТАР ТЕЛЕНЕН ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит
Корпуста эзлөү
Татарча сөйләгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат



Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatacorpus@gmail.com

Сүз #1   **Сүзформа:**  
Лемма:  
Алдагы сүзгә кадәр ара: 1  **тигез / алып** 1  **кадәр** 

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренәң артыннан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула
- <...> ! Ideophone ! Арау илгәтә ! Эшкәрәткәчәтә ! Исе сөзә

Эксперименталь фонетика лабораториясе

Тәрәзәнең өске сул чатында хәзерге вакытта карала торган сүзнең номеры күрсәтелә. Аның уң ягында ике кнопка урнаштырылган, алар арттагы яки алдагы сүзгә тиз күчү функциясен башкаралар. Күчеш вакытында әлеге биттәге мәгълүматның автоматик рәвештә саклануын истә тотарга кирәк!

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Сүз #1

Сүзформа:

Лемма:

Алдагы сүзгә кадәр ара: 1 тигез / алып 1 кадәр

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренең артынан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула

Ok Ябу

Эксперименталь фонетика лабораториясе



Сүзформа, лемма һәм араны күрсәтү аланчыклары рәсемдә кызыл төс белән бирелгән:

ТАТАР ТЕЛЕНЕН ЯЗМА КОРПУСЫ

TAT РУС ENG





Баш бит
Корпуста эзлөү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.
tatorpus@gmail.com

Сүз #1  

Сүзформа:

Лемма:

Алдагы сүзгә кадәр ара: 1  **тигез / алып** 1  **кадәр**  

Грамматик билгеләр

! Part of speech categories = First-level tags










- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренәң артыннан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула

Эксперименталь фонетика лабораториясе

Тәрәзәнең төп өлешендә грамматик билгеләрне сайлап алу исемлеге категорияләргә бүленеп урнаштырылган. Теге яки бу тегны күрсәтү өчен шул юлны “тамгаларга” кирәк.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Сүз #1   Сүзформа:  
Лемма:  
Алдагы сүзгә кадәр ара: 1  тигез / алып 1  кадәр 

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренә артынан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула
- <id> ! Ideophone ! Арау чатма ! Эпиграммалар, шул сүзләр

Эксперименталь фонетика лабораториясе

Баш бит
Корпуса эзлөү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган макаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com







Өске уң чатта һәм аста мәгълүматны саклап (яшел) яки сакламыйча (кызыл) тәрәзәне ябу теймәләре урнашкан.

ТАТАР ТЕЛЕНЕН ЯЗМА КОРПУСЫ

TAT РУС ENG

Баш бит
Корпуста эзләү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

Сүз #1   **Сүзформа:**  
Лемма:
Алдагы сүзгә кадәр ара: 1  **тигез / алып** 1  **кадәр**

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренәң артыннан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула
- <idc> ! Ideophone ! Арау чаттема ! Эпиграммалар, шулар





Ok **Ябу**



Эксперименталь фонетика лабораториясе




Тагын бер мисал буларак “алма” (исем сүз төркеме) сүзформасын эзләүне карыйк. Моның өчен сүзформа аланчыгында “алма” дип язабыз, ә теглар исемлегендә исем сүз төркемен сайлап алабыз.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG

Сүз #1   Сүзформа: алма  

Лемма:  

Алдагы сүзгә кадәр ара: 1  тигез / алып 1  кадәр 

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренәң артынан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула
- <idc> ! Ideophone ! Аңаң картма ! Эңгәләткәчтәшчә сүз



Ok Ябу



Эксперименталь фонетика лабораториясе



Өлеге сүзформаны эзләнең икенче юлы буларак сүзформаны да, лемманы да “алма” дип күрсәтү тора. Бу очракта система “алма” фигылен кире кагачак, чөнки аның леммасы -- “ал”.




ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG

СҮЗ #1  

Сүзформа: алма  



Лемма: алма  

Алдагы сүзгә кадәр ара: 1  тигез / алып 1  кадәр 

Грамматик билгеләр

! Part of speech categories = First-level tags

- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләренең артынан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз
- <cnjadv> ! Adverbial conjunction ! Рәвеш-теркәгеч ! Наречие-союз
- <ij> ! Interjection ! Ымлык ! Междометие
- <abbr> ! Abbreviation ! Кыскартылма ! Аббревиатура
- <cop> ! Copula ! Копула ! Копула
- <idc> ! Ideophone ! Аңа нәрсә ! Эпиграммалар, шул сүзләр

Эксперименталь фонетика лабораториясе



Түбөндөгө рәсемдә “ас*” хәрефләрәнә башланган бөтен фигыльләрне (<v>) табу карала:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

Баш бит
Корпуса эзләү
Татарча сөйлөгеч
Кулланма
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында
Файдалы мәгълүмат

Авторлар:
Сайдунов М.Р.
Ибраһимов Т.И.
tatcorpus@gmail.com

Сүз #1  

Сүзформа:

Лемма:

Сүзнең башы:

Сүзнең азагы:

Алдагы сүзгә кадәр ара: тигез / алып кадәр

Грамматик билгеләр

! Part of speech categories = First-level tags

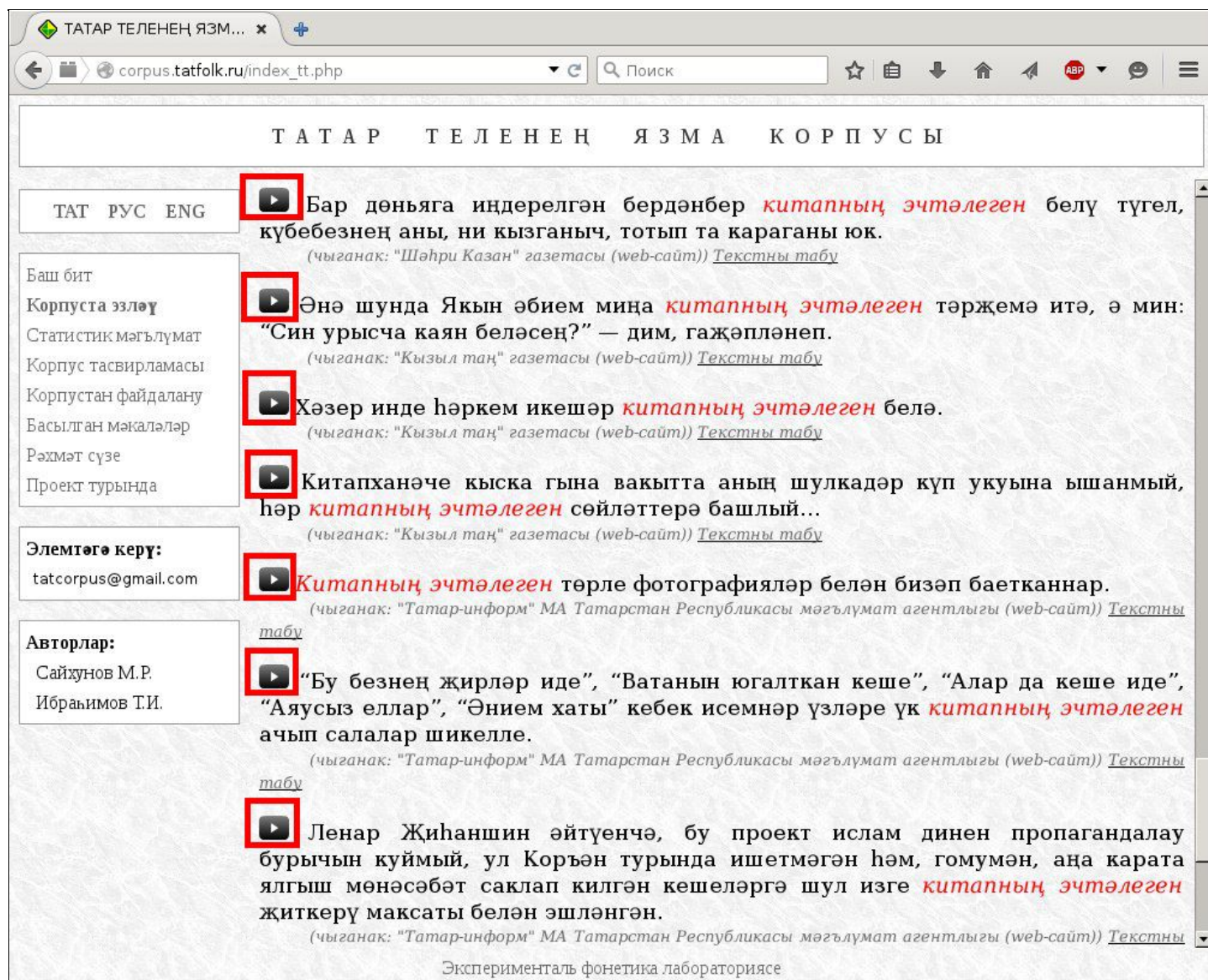
- <n> ! Noun ! Исем ! Существительное
- <np> ! Proper noun ! Ялгызлык исем ! Имя собственное
- <adj> ! Adjective ! Сыйфат ! Прилагательное
- <num> ! Numeral ! Сан ! Числительное
- <prn> ! Pronoun ! Алмашлык ! Местоимение
- <det> ! Determiner ! Детерминатив ! Детерминатив
- <v> ! Verb ! Фигыль ! Глагол**
- <vaux> ! Auxiliary verb ! Ярдәмче фигыль ! Вспомогательный глагол
- <adv> ! Adverb ! Рәвеш ! Наречие
- <post> ! Postposition ! Бәйләк ! Послелог
- <postadv> ! Postadverb ! Пострәвеш ! Посленаречие (Ачыклаган сүзләрәнәң артыннан килүче рәвешләр (кисәкчәләр) !# гына, ук)
- <cnjcoo> ! Coordinating conjunction ! Тезүче теркәгеч ! Сочинительный союз
- <cnjsub> ! Subordinating conjunction ! Ияртүче теркәгеч ! Подчинительный союз

Ok Ябу

Эксперименталь фонетика лабораториясе

Мисаллар буларак китерелгән жөмлөләрне тыңлап карау

Татар теленең Язма корпусынан файдаланучыларга мисал буларак китерелгән жөмлөләрне тыңлап карау мөмкинлеге эшләнде. Аерым жөмлөнә тыңлау өчен шөхес аның сул ягында урнашкан кнопкага чиртергә тиеш була.



The screenshot shows the website 'ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ' (Tatar Written Corpus). The browser address bar shows 'corpus.tatfolk.ru/index_tt.php'. The page has a navigation menu with 'ТАТ РУС ENG' and a sidebar with links like 'Баш бит', 'Корпуста эзләү', 'Статистик мәгълүмат', 'Корпус тасвирламасы', 'Корпустан файдалану', 'Басылган мәкаләләр', 'Рәхмәт сүзе', and 'Проект турында'. Below the sidebar, there are sections for 'Элемтөгә керү:' (Email: tatcorpus@gmail.com) and 'Авторлар:' (Authors: Сайхунов М.Р., Ибраһимов Т.И.). The main content area lists several text samples, each with a red box around a play button icon. The samples are:

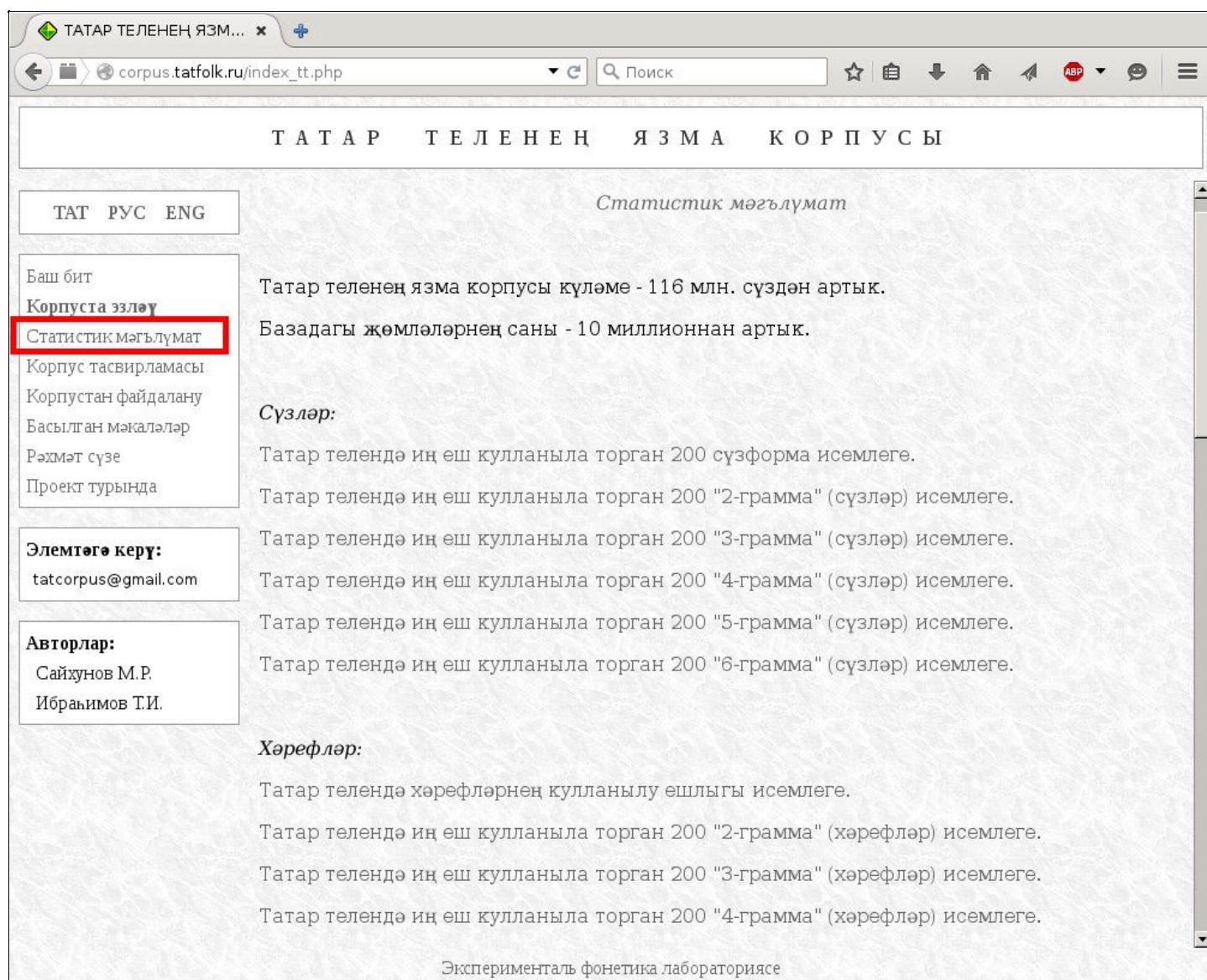
- 1. **Бар дөньяга индерелгән бердәнбер *китапның эчтәлеген* белү түгел, күбебезнең аны, ни кызганыч, тотып та караганы юк.** (чыганақ: "Шәһри Казан" газетасы (web-сайт)) [Текстны табу](#)
- 2. **Әнә шунда Якын әбием миңа *китапның эчтәлеген* тәржемә итә, ә мин: "Син урысча каян беләсең?" — дим, гажәпләнәп.** (чыганақ: "Кызыл таң" газетасы (web-сайт)) [Текстны табу](#)
- 3. **Хәзер инде һәркем икешәр *китапның эчтәлеген* белә.** (чыганақ: "Кызыл таң" газетасы (web-сайт)) [Текстны табу](#)
- 4. **Китапханәче кыска гына вакытта аның шулкадәр күп укуына ышанмый, һәр *китапның эчтәлеген* сөйләттерә башлый...** (чыганақ: "Кызыл таң" газетасы (web-сайт)) [Текстны табу](#)
- 5. ***Китапның эчтәлеген* төрле фотографияләр белән бизәп баетканнар.** (чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)
- 6. **"Бу безнең жирләр иде", "Ватанын югалткан кеше", "Алар да кеше иде", "Аяусыз еллар", "Әнием хаты" кебек исемнәр үзләре үк *китапның эчтәлеген* ачып салалар шикелле.** (чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)
- 7. **Ленар Жиһаншин әйтүенчә, бу проект ислам динен пропагандалау бурычын куймый, ул Корьән турында ишетмәгән һәм, гомумән, аңа карата ялгыш мөнәсәбәт саклап килгән кешеләргә шул изге *китапның эчтәлеген* житкерү максаты белән эшләнгән.** (чыганақ: "Татар-информ" МА Татарстан Республикасы мәгълүмат агентлыгы (web-сайт)) [Текстны табу](#)

Эксперименталь фонетика лабораториясе

Өстәмә статистик мәгълүмат!

Корпусның авторлары эш барышында жыелган төрле статистик мәгълүматны сайтның әлеге бүлегендә урнаштыралар. Бүгенге көндә шундыйларга түбәндәгеләр карый:

- Татар телендә иң еш кулланыла торган сүзформалар исемлеге.
- 2, 3, 4, 5 һәм 6 сүздән торган зур ешлыклы тезмәләр исемлеге.
- Татар телендә хәрәфләрнең һәм аларның төрле комбинацияләре нигезендә төзелгән ешлыклар исемлеге.



The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content area is titled "Статистик мәгълүмат" and contains the following text:

Татар теленең язма корпусы күләме - 116 млн. сүздән артык.
Базадагы жөмлөләрнең саны - 10 миллионнан артык.

Сүзләр:
Татар телендә иң еш кулланыла торган 200 сүзформа исемлеге.
Татар телендә иң еш кулланыла торган 200 "2-грамма" (сүзләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "3-грамма" (сүзләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "4-грамма" (сүзләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "5-грамма" (сүзләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "6-грамма" (сүзләр) исемлеге.

Хәрәфләр:
Татар телендә хәрәфләрнең кулланылу ешлыгы исемлеге.
Татар телендә иң еш кулланыла торган 200 "2-грамма" (хәрәфләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "3-грамма" (хәрәфләр) исемлеге.
Татар телендә иң еш кулланыла торган 200 "4-грамма" (хәрәфләр) исемлеге.

Эксперименталь фонетика лабораториясе

On the left sidebar, the "Статистик мәгълүмат" link is highlighted with a red box. Other links include "Баш бит", "Корпуста эзлө", "Корпус тасвирламасы", "Корпустан файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтәгә керү: tatcorpus@gmail.com", and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И."

- Хәрефләрнең сүз башында һәм сүз азагында кулланылу ешлыгы исемлекләре.
- Фонемаларның һәм аларның комбинацияләренең сүз эчендә һәм ритмик төркем кысаларында кулланылу ешлыгын гәүдәләндергән исемлекләр.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT RUS ENG

Баш бит
Корпуста эзләү
 Статистик мәгълүмат
 Корпус тасвирламасы
 Корпустаң файдалану
 Басылган мәкаләләр
 Рәхмәт сүзе
 Проект турында

Элемтәгә керү:
 tatcorpus@gmail.com

Авторлар:
 Сайхуннов М.Р.
 Ибраһимов Т.И.

Хәрефләр (сүз башында):

Татар телендә иң еш кулланыла торган 200 "2-грамма" (хәрефләр, сүз башында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "3-грамма" (хәрефләр, сүз башында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "4-грамма" (хәрефләр, сүз башында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "5-грамма" (хәрефләр, сүз башында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "6-грамма" (хәрефләр, сүз башында) исемлегә.

Хәрефләр (сүз азагында):

Татар телендә иң еш кулланыла торган 200 "2-грамма" (хәрефләр, сүз азагында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "3-грамма" (хәрефләр, сүз азагында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "4-грамма" (хәрефләр, сүз азагында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "5-грамма" (хәрефләр, сүз азагында) исемлегә.

Татар телендә иң еш кулланыла торган 200 "6-грамма" (хәрефләр, сүз азагында) исемлегә.

Эксперименталь фонетика лабораториясе

Монда татар теленең иң еш кулланыла торган 5000 сүзформасы китерелә:

The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content is a table listing 45 words and their frequencies. The table is as follows:

1	һәм	1630726
2	белән	1249114
3	да	1204431
4	дә	910669
5	бу	746627
6	ул	727682
7	дип	638244
8	өчен	603250
9	бер	508703
10	иде	417602
11	аның	353914
12	буенча	341434
13	ә	340479
14	татар	309106
15	гәнә	289810
16	бик	286219
17	инде	283772
18	шул	269870
19	бар	265782
20	түгел	264347
21	гына	264131
22	шулай	257466
23	Татарстан	255036
24	үз	251835
25	ук	251452
26	турында	242908
27	алар	242730
28	булган	242209
29	мин	241080
30	булып	239581
31	ел	232660
32	зур	225474
33	кеше	224115
34	алып	218246
35	иң	217754
36	яңа	213919
37	итеп	207815
38	дәүләт	207369
39	соң	204223
40	кирәк	198707
41	әлеге	198508
42	беренче	196135
43	торган	194072
44	күп	192507
45	то	191750

Below the table, the text "Эксперименталь фонетика лабораториясе" is visible. The left sidebar contains navigation links for "ТАТ РУС ENG", "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустан файдалану", "Басылган мәкаләләр", "Рәхмәт сүзе", "Проект турында", "Элемтөгә керү: tatcorpus@gmail.com", and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И."

Беренче баганда тәртип саны, икенчесендә - сүзформа, өчөнкесендә - Корпуста аларның кулланылу ешлыгы бирелә.

Ә монда ике хәрефтән торган тезмәләрнең ешлыгы күрсәтелгән:

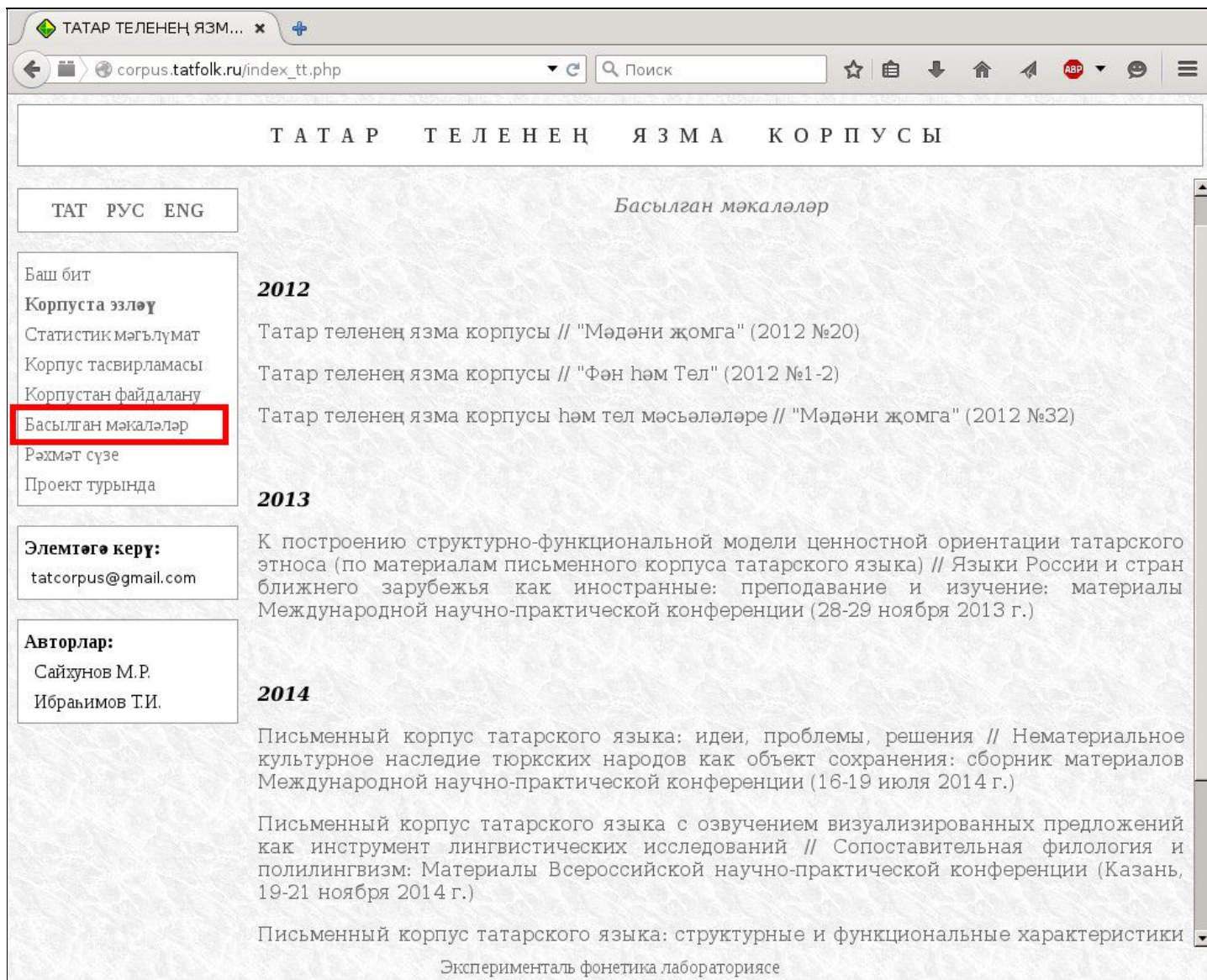
The screenshot shows a web browser window with the URL `corpus.tatfolk.ru/index_tt.php`. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main content is a list of two-letter words in Tatar, numbered 1 to 45, with their corresponding frequencies and the words themselves. The browser's address bar shows "Поиск" and the page navigation indicates "Страница: 1 из 4".

№	Сүз	Эш
1	ар	13487484
2	ла	10927856
3	ан	10537847
4	лә	9389724
5	ен	8426490
6	әр	7742249
7	ал	7508939
8	га	7476037
9	та	7226503
10	ын	6947006
11	да	6721316
12	ел	6716396
13	ка	6538509
14	ер	6512923
15	ән	6420902
16	ны	5772159
17	ре	5497572
18	нд	5470991
19	ры	5292593
20	дә	5203472
21	на	5122055
22	ле	5101267
23	ра	5056469
24	бе	4975734
25	лы	4812588
26	ул	4644417
27	ат	4600942
28	те	4458442
29	ба	4039938
30	не	3887440
31	ма	3817973
32	ты	3731737
33	ге	3708165
34	гә	3700021
35	бу	3649670
36	гы	3507257
37	тә	3492769
38	ин	3456093
39	ак	3448221
40	нә	3393397
41	мә	3337527
42	ки	3275056
43	ит	3209149
44	ыр	3175581
45	ог	3146247

Эксперименталь фонетика лабораториясе

Басылган мәкаләләр

Сайтның “Басылган мәкаләләр” бүлегендә Корпус авторлары тарафыннан язылган мәкаләләр урнаштырыла:



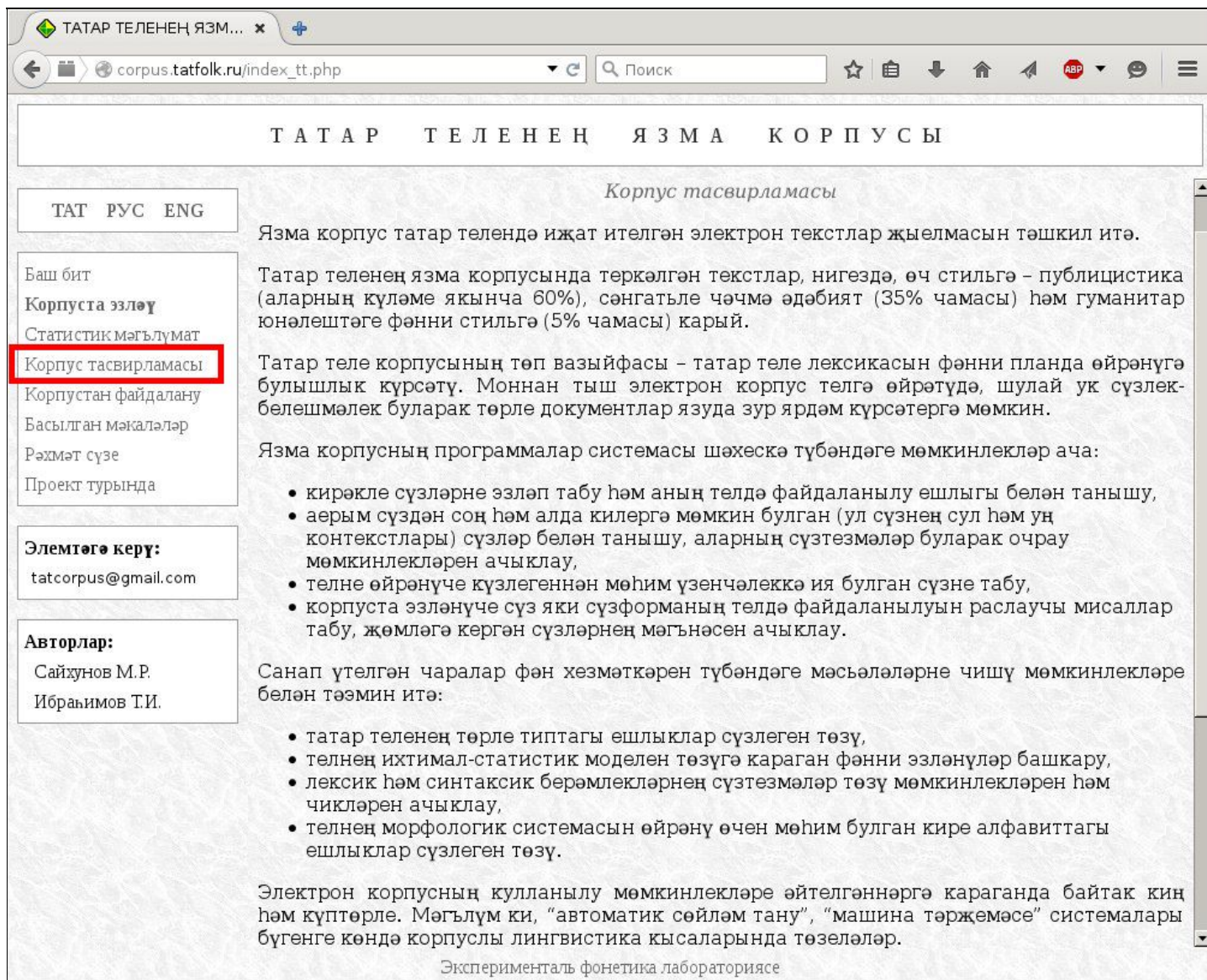
The screenshot shows a web browser window with the URL corpus.tatfolk.ru/index_tt.php. The page title is "ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ". The main heading is "Басылган мәкаләләр". On the left, there is a navigation menu with options: "ТАГ РУС ENG", "Баш бит", "Корпуста эзләү", "Статистик мәгълүмат", "Корпус тасвирламасы", "Корпустаң файдалану", "Басылган мәкаләләр" (highlighted with a red box), "Рәхмәт сүзе", and "Проект турында". Below the menu, there are contact details: "Элемтәгә керү: tatcorpus@gmail.com" and "Авторлар: Сайхунов М.Р., Ибраһимов Т.И.". The main content area lists articles by year:

- 2012**
 - Татар теленең язма корпусы // "Мәдәни жомга" (2012 №20)
 - Татар теленең язма корпусы // "Фән һәм Тел" (2012 №1-2)
 - Татар теленең язма корпусы һәм тел мәсьәләләре // "Мәдәни жомга" (2012 №32)
- 2013**
 - К построению структурно-функциональной модели ценностной ориентации татарского этноса (по материалам письменного корпуса татарского языка) // Языки России и стран ближнего зарубежья как иностранные: преподавание и изучение: материалы Международной научно-практической конференции (28-29 ноября 2013 г.)
- 2014**
 - Письменный корпус татарского языка: идеи, проблемы, решения // Нематериальное культурное наследие тюркских народов как объект сохранения: сборник материалов Международной научно-практической конференции (16-19 июля 2014 г.)
 - Письменный корпус татарского языка с озвучением визуализированных предложений как инструмент лингвистических исследований // Сопоставительная филология и полилингвизм: Материалы Всероссийской научно-практической конференции (Казань, 19-21 ноября 2014 г.)
 - Письменный корпус татарского языка: структурные и функциональные характеристики

Эксперименталь фонетика лабораториясе

Корпус тасвирламасы

Сайтның “Корпус тасвирламасы” битендә Татар теле язма корпусының структурасы һәм функциональге турында кыскача информация тәкъдим ителә:



ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

Корпус тасвирламасы

ТАГ РУС ENG

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустан файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтәгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхуннов М.Р.
Ибраһимов Т.И.

Язма корпус татар телендә иҗат ителгән электрон текстлар җыелмасын тәшкил итә.

Татар теленең язма корпусында теркәлгән текстлар, нигездә, өч стильгә – публицистика (аларның күләме якынча 60%), сәнгатьле чөмә әдәбият (35% чамасы) һәм гуманитар юнәлештәге фәнни стильгә (5% чамасы) карый.

Татар теле корпусының төп вазыйфасы – татар теле лексикасын фәнни планда өйрөнүгә булышлык күрсәтү. Моннан тыш электрон корпус телгә өйрәтүдә, шулай ук сүзлек-белешмәлек буларак төрле документлар язуда зур ярдәм күрсәтергә мөмкин.

Язма корпусның программалар системасы шөхескә түбәндөгә мөмкинлекләр ача:

- кирәкле сүзләрне эзләп табу һәм аның телдә файдаланылу ешлыгы белән танышу,
- аерым сүздән соң һәм алда килергә мөмкин булган (ул сүзнең сул һәм уң контекстлары) сүзләр белән танышу, аларның сүзтөзмәләр буларак очрау мөмкинлекләрен ачыклау,
- телне өйрөнүче күзлегеннән мөһим үзенчәлеккә ия булган сүзне табу,
- корпуста эзләнүче сүз яки сүзформаның телдә файдаланылуын раслаучы мисаллар табу, җөмлөгә кергән сүзләрнең мәгънәсен ачыклау.

Санап үтелгән чаралар фән хезмәткәрен түбәндөгә мәсьәләләрне чишү мөмкинлекләре белән тәэмин итә:

- татар теленең төрле типтагы ешлыктар сүзлеген төзү,
- телнең ихтимал-статистик моделен төзүгә караган фәнни эзләнүләр башкару,
- лексик һәм синтаксик берәмлекләрнең сүзтөзмәләр төзү мөмкинлекләрен һәм чикләрен ачыклау,
- телнең морфологик системасын өйрөнү өчен мөһим булган кире алфавиттагы ешлыктар сүзлеген төзү.

Электрон корпусның кулланылу мөмкинлекләре әйтелгәннәргә караганда байтак кин һәм күптөрле. Мәгълүм ки, “автоматик сөйләм тану”, “машина тәрҗемәсе” системалары бүгенге көндә корпуслы лингвистика кысаларында төзеләләр.

Эксперименталь фонетика лабораториясе

Корпустаң файдалану

Сайтның “Корпустаң файдалану” бүлегендә Корпустаң файдалануның хокукий нормалары белән һәм Корпуска һәм аңардагы материалларга ссылгалар һәм цитаталар ясау турында төхъдимнәр белән танышырга мөмкин.

Эшегездә Корпусны кулланган очракта, авторлар сездән мөкалэләрегездә Корпуска һәм корпус авторлары язган мөкалэләргә сылтама ясарга онытмавыгызны сорыйлар.

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАТ РУС ENG баш биткә

Корпустаң файдалану

Интернет чөлтөрөнә <http://corpus.tatfolk.ru> адресы белән урнаштырылган, татар теленә язма корпусында файдаланылган барлык интеллектуаль эш нәтижәләре (1274 ГК РФ статьясына тугрылыкта) коммерциягә кагылышы булмаган, бары тик фәнни-тикшерү һәм уку-укыту максаты белән куллануны күздә тотта.

Аларны шулай ук уку да (карап чыгу да), күчәрәп алу да, башка формаларда файдалану да рөхсәт ителми. Аларны бары тик кайсыдыр төл күренешен мисаллар белән жөпләүдә, эзләү режимында мисаллар (цитаталар) чыганагы буларак файдаланырга мөмкин.

Татар теленә язма корпусы ярдәмендә табылган цитаталар китерелгәндә "Татар теленә язма корпусы"на һәм мисал чыганагына сылтау ясау зарур.

Татар теле корпусының оффлайн версиясә чикле күләмдә берәз соңрак китереләчөк!

Эксперименталь фонетика лабораториясә

Баш бит
Корпуста эзләү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мөкалэләр
Рәжмәт сүзе
Проект турында

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхунов М.Р.
Ибраһимов Т.И.

Проект турында

Сайтның “Проект турында” бүлегендә Корпусны ясау тарихы һәм аны төп эшләүчеләр турында мәгълүмат бирелә:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

Проект турында

ТАГ РУС ENG

Баш бит
Корпуста эзлөү
Статистик мәгълүмат
Корпус тасвирламасы
Корпустаң файдалану
Басылган мәкаләләр
Рәхмәт сүзе
Проект турында

Элемтөгә керү:
tatcorp.us@gmail.com

Авторлар:
Сайхуннов М.Р.
Ибраһимов Т.И.

Татар теленең язма корпусын төзүгә караган эшләрнең башлануы 2010 елга карый. Казан федераль университетның Гамәли лингвистика һәм эксперименталь фонетика лабораториясендә узган фәнни юнәлеш сайлауга караган фикер алышуларда түбөндөгә ике тема тәкъдим ителә:

- татарча текстларны башка бер тугандаш телгә һәм, киресенчә, кайсыдыр кардәш телдәгә текстларны татарчага машина ярдәмендә тәржемә итү системасын булдыру,
- аерым темага караган татарча сөйләмне танучы (язма формага күчерүче) автоматлаштырылган система төзү.

Әдәбият белән танышулар дәвамында бүгенгә автоматик машина тәржемәсе һәм сөйләм тану системаларының эше милли тел корпусларын – “гипотеза – ачыклау” алымын файдалануга корылганлыгы мәгълүм булды. Бу хәл безне татар теленең милли корпусын төзүгә ныклап алынырга мәжбүр итте.

Язма корпус, нигездә, web-ресуслар материалында төзелде. Китерелгән мисалларның сайт адреслары аркылы (һәр сүзформаның кулланышы жөмлөләр ярдәмендә тасвирлана) корпуста файдаланылган сайтларга карата тулы мәгълүмалар алырга мөмкин.

Татар теленең язма корпусына теркәлүче текстларга башлангыч эшкәртү ясалды – html тегыларыннан чистарту, чит телдә язылган жөмлөләрдән азат итү, utf-8 гә күчерү, жөмлө чикләрен автоматик рәвештә тамгалау һ.б. эшләр үтәлдә.

Хәзергә вакытта материал жыю, электрон корпуска теркәү өчен аларны эшкәртү дәвам итә. Кайбер язучыларның, фән хезмәткәрләренең татар теле корпусы төзелүен белеп, үз китапларының электрон версиясен безгә тәкъдим итү мисаллары да бар. Бу материалларның күләме 5 – 6 миллион сүзгә житү белән алар текстлар корпусына теркәлчәк. Бер уңайдан, язма корпусның функцияль мөмкинлекләре дә баетыла.

Өйтергә кирәк, **Татар теленең язма корпусы** – ул татар тел дөньясын бер тәртиптә тасвирлаучы гаять зур сүзлек – бөлешмә тәшкил итә.

Татар теле корпусыннан файдалану түләүсез.

Эксперименталь фонетика лабораториясе

Рәхмәт сүзе

“**Рәхмәт сүзе**” битендә, исемнен күренгәнчә, Корпусның авторлары бу эштә зур ярдәм күрсәткән кешеләргә һәм оешмаларга рәхмәт сүзен әйтәләр:

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

ТАГ РУС ENG баш биткә

Зур рәхмәтлебез!

Рәхмәтебезне һәм тирән ихтирамыбызны белдерәбез:

- Республиканың традицион мәдәниятны үстерү үзәгенә һәм аның директоры Завгарова Фәнзилә Хәким кызына www.tatfolk.ru хостингысында сайтка урын һәм корпуска электрон текстлар бирүләре өчен.
- Турку университетының Фин-угор телләре кафедрасына һәм шәхсән Йорма Луутоненга корпуска дип бирелгән текстлар һәм файдалы киңәшләре өчен.
- "Фән һәм Тел" исемле фәнни-мәгълүмати журнал редакциясенә, аның редакторы Шакиржанов Рәшит Әгълә улына һәм шулай ук Шакиржанов Фәриткә күпсанлы текстлар һәм проектка ярдәмнәре өчен.
- Хусаинов Р.Р. ("GDC" компаниясе) - корпусны эшләү дәвамындагы бик зур ярдәмә өчен.
- Чугунов А.Н. ("RX5" компаниясе) - мәгълүмат базаларының структурасы һәм эзләнү системасын тизәйтү өлкәсендәгә мөһим киңәшләре өчен.

БЕЗГӘ ҮЗ ХЕЗМӨТЛӨРЕН ҺӘМ ЯЗМАЛАРЫН ЖИБӨРГӨН БӨТЕН АВТОРЛАРГА!!!

Эксперименталь фонетика лабораториясе

Элемтөгә керү өчен мәгълүмат

Корпус авторларының е-mail'ы: tatcorpus@gmail.com

Татар теленең язма корпусын куллануга һәм үстерүгә бәйлә сорауларыгыз һәм тәкъдимнәрегез белән Безгә мөрәжәгать итүегезне сорыйбыз!

ТАТАР ТЕЛЕНЕҢ ЯЗМА КОРПУСЫ

TAT РУС ENG

Гомуми мәгълүмат

Бу сайтта **замани татар теленең язма корпусы** урнаштырылган. Язма корпус бүгенге татар телен электрон формада гәүдәләндерә. Хәзерге вакытта татар теле корпусының күләме **116 миллионнан артык сүз** тәшкил итә, төрле сүзформаларның саны 1 миллионнан артык. Язма корпус татар теленең системасы, хәле, алдагы көне белән кызыксынучы шәхесләрне күздә тотып төзелә. Ул татар телен корпуслы лингвистика кысаларында өйрәнүче белгечләр өчен зарури чыганақ.

Элемтөгә керү:
tatcorpus@gmail.com

Авторлар:
Сайхуннов М.Р.
Ибраһимов Т.И.

Проект яңалыклары

18.04.2015 - Корпуста **калып (сүз азагы)** нигезендә эзләү системасы гамәлгә кертелде.

29.03.2015 - Уң, сул һәм семантик контекстларны күрсәтү санының чиге 100 дән 10 000 гә кадәр житкерелде. Аларны таблица буларак күрү өчен "Барысын да карау" төймәсен кулланыгыз.

26.03.2015 - Хәзер Корпуска яңа **corpus.tatar** адресы аша да керергә мөмкин. Шул ук вакытта иске **corpus.tatfolk.ru** адресы да эшләчәк

14.03.2015 - Корпуста **калып (сүз башы)** нигезендә эзләү системасы гамәлгә кертелде.

12.10.2014 - Тәкъдим ителгән **жөмлөләрне тыңлату** мөмкинлегә булдырылды (жөмләнән сул ягында урнашкан махсус төймәгә басып).

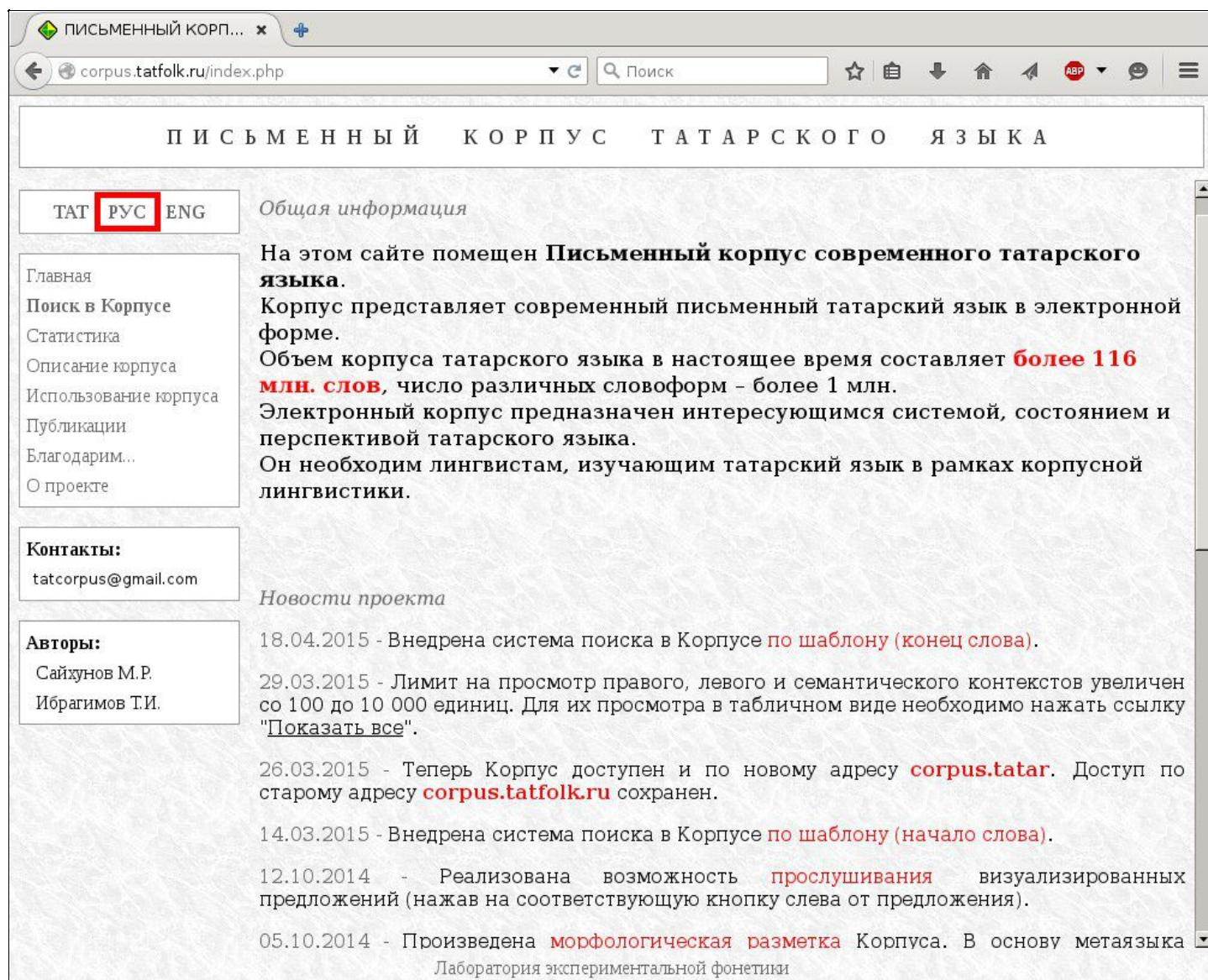
05.10.2014 - Корпусны **морфологик тамгалау** башкарылды. Грамматик тамгалар метателе нигезенә Апертиум халыкара проекты тарафыннан төрки телләр өчен эшләнгән «тег»лар системасы алынды.

Эксперименталь фонетика лабораториясе

Башка телләрҙә

Кулланучыларыбызга уңайлырак булсын өчен Татар теле язма корпусы сайтының бөтен бүлекләре дә бүгенге көндә өч телдә тәкъдим ителә: русча, татарча һәм инглизчә. Башка теллөргә тәржемә итү эше дә бара!

Интерфейсның телен сайлау өчен экранның сул ягында төп меню өстендәге тиешле сыылкага басарга кирәк (бу очракта рус телендә):



ПИСЬМЕННЫЙ КОРПУС ТАТАРСКОГО ЯЗЫКА

TAT **РУС** ENG

Общая информация

На этом сайте помещен **Письменный корпус современного татарского языка**. Корпус представляет современный письменный татарский язык в электронной форме. Объем корпуса татарского языка в настоящее время составляет **более 116 млн. слов**, число различных словоформ – более 1 млн. Электронный корпус предназначен интересующимся системой, состоянием и перспективой татарского языка. Он необходим лингвистам, изучающим татарский язык в рамках корпусной лингвистики.

Новости проекта

18.04.2015 - Внедрена система поиска в Корпусе [по шаблону \(конец слова\)](#).

29.03.2015 - Лимит на просмотр правого, левого и семантического контекстов увеличен со 100 до 10 000 единиц. Для их просмотра в табличном виде необходимо нажать ссылку ["Показать все"](#).

26.03.2015 - Теперь Корпус доступен и по новому адресу [corpus.tatar](#). Доступ по старому адресу [corpus.tatfolk.ru](#) сохранен.

14.03.2015 - Внедрена система поиска в Корпусе [по шаблону \(начало слова\)](#).

12.10.2014 - Реализована возможность [прослушивания](#) визуализированных предложений (нажав на соответствующую кнопку слева от предложения).

05.10.2014 - Произведена [морфологическая разметка](#) Корпуса. В основу метаязыка

Лаборатория экспериментальной фонетики

Главная
Поиск в Корпусе
Статистика
Описание корпуса
Использование корпуса
Публикации
Благодарим...
О проекте

Контакты:
tatcorpus@gmail.com

Авторы:
Сайхунов М.Р.
Ибрагимов Т.И.

Ә бу исә Корпус сайтының англиз телендәге бите:

CORPUS OF WRITTEN TATAR

TAT РУС **ENG**

Main
Search the Corpus
Statistics
Description
Using the Corpus
Publications
Acknowledgements
About the project

Contact address:
tatcorpus@gmail.com

Developers:
Saykhunov M.R.
Ibragimov T.I.

General Information

This website contains a **Text Corpus of the modern Tatar language** consisting of **over 116 million word occurrences**.
The corpus represents modern written Tatar language in electronic form. The total count of different word forms in the Tatar corpus is more than 1 mln. This collection of Tatar texts in electronic form is intended for the use of those interested in the structure, present condition and prospects of the Tatar language.
The Corpus of Written Tatar language is indispensable for everyone who wants to study Tatar by the methods of corpus linguistics.

Project news

18.04.2015 - **Template (the end of the word)** based search system in the Corpus is implemented.

29.03.2015 - Limit for viewing right, left and semantic contexts is increased from 100 to 10 000 units. To view them in a table format you should click on the link "[Show all](#)".

26.03.2015 - Now The Corpus is also available at the new address **corpus.tatar**. At the same time you can use the old address **corpus.tatfolk.ru**.

14.03.2015 - **Template (the beginning of the word)** based search system in the Corpus is implemented.

12.10.2014 - Implementation of a system for **listening** of visualized sentences (by clicking on the appropriate button on the left of sentences)

05.10.2014 - The **morphological marking** of the Corpus is made. The meta-language of

Laboratory of Experimental Phonetics