

**ТЮРКСКИЙ МИР КАЗАХСТАНА: ИССЛЕДОВАНИЕ ЯЗЫКОВ И
СОЗДАНИЕ БАЗЫ ДАННЫХ
TURKIC WORLD OF KAZAKHSTAN: RESEARCH ON THE
LANGUAGES AND DATA BASE CREATION**

Невская Ирина Анатольевна*
Тажибаева Сауле Жаксылыкбаевна**
Шаймердинова Нурила Габбасовна***
Тусупов Джамалбек Алиаскарович****
Schönig Claus *****

АННОТАЦИЯ:

Тюркский лингвистический ландшафт Казахстана представляет интересный объект для социолингвистического и лингвокультурологического исследования. Лингвистический и этнографический материал, собранный во время полевых исследований (2014-2016), данные социолингвистического опроса (2013-2016) представляют собой ценный материал для тюркологов и намечают новые перспективы тюркологических и социолингвистических исследований в РК.

КЛЮЧЕВЫЕ СЛОВА:

тюркские языки Казахстана, документация, социолингвистическая ситуация.

В Казахстане проживает 25 тюркских этноса, говорящих на разных тюркских языках кыпчакской, огузской и карлукской групп.

Исследователи выделяют экзогенные и эндогенные языки диаспоры Казахстана [Сулейменова 2007: 26]. Большинство казахстанских тюркских языков относятся к группе экзогенных языков. Это - языки, которые за пределами Казахстана имеют свою территорию, представляют численное большинство носителей, т.е. являются мажоритарными [Trudgill 2003: 43; Сулейменова 2007: 26]. Нередко оба эти критерия вместе не могут быть

* д.ф.н., профессор Frankfurt University/Institute of Empirical Linguistics; г.н.с. Институт филологии СО РАН E-mail: nevskaya@em.uni-frankfurt.de

** д.ф.н., профессор ЕНУ им. Л.Н.Гумлева. E-mail: tazhibaeva_szh@enu.kz

*** д.ф.н., профессор ЕНУ им. Л.Н.Гумлева. E-mail: shaimerdinova_ng@enu.kz

**** д.ф.-м.н., профессор ЕНУ им. .Н.Гумлева. E-mail: tusupov_da@enu.kz

***** доктор, профессор Берлинский Свободный университет E-mail: clcs@gmx.de

применены, так как даже при наличии собственных административных территорий представители титульных этносов часто не составляют большинство их населения. В таком случае мы руководствуемся только первым критерием, - т.е. наличием титульной территории, - при отнесении тех или иных этносов к экзогенным. Среди тюркских языков Казахстана к экзогенным относятся следующие языки: узбекский, уйгурский, азербайджанский, татарский, башкирский, чувашский, киргизский, каракалпакский, туркменский, ногайский, кумыкский, карачаевский, балкарский, гагаузский, алтайский, хакасский, якутский, турецкий (ахыска). В таблице №1 представлено количество носителей экзогенных языков в соответствии с данными переписей населения за период 1970-2009 г.г. www.stat.gov.kz:

Таблица №1.

Количественное соотношение носителей тюркских экзогенных языков

	1970	1979	1989	1999	2009
Общее кол-во населения	13.026.274	14.709.508	16.222.324	14.981.281	16.009.597
1. Узбеки	216.258	262.960	330.417	370.765	456.997
2. Уйгуры	120.622	147.676	181.155	210.377	224.713
3. Татары	286.878	314.065	322.338	249.052	204.229
4. Башкиры	21.500	32.577	41.060	23.247	17.263
5. Каракалпаки	463	620	1.357	1.497	2.828
6. Киргизы	9.612	9.352	13.718	10.925	23.274
7. Азербайджанцы	57.607	73.240	88.887	78.325	85.292
8. Туркмены	3.265	2.241	3.716	1.733	2.234
9. Гагаузы	772	752	953	678	493
10. Кумыки	554	873	1.700	643	481
11. Ногайцы	155	236	539	350	276
12. Тувинцы	85	182	129	35	37
13. Чуваши	22.871	22.310	21.717	11.864	7.301
14. Якуты	175	438	303	115	119
15. Алтайцы	575	630	675	462	221
16. Хакасы	401	475	575	355	223
17. Карачайцы	2.447	2.082	2.038	1.400	995
18. Балкарцы	2.714	2.258	2.926	2.079	1.798
19. Турки	18.377	25.718	49.219	78.711	97.015

За период последней переписи населения (1999-2009 г.г.) среди представителей экзогенных языков произошел значительный рост населения. Более чем в два раза увеличилось количество представителей киргизской диаспоры (113%), на 89 % - каракалпаков, туркмен на 29 %, узбеков, турок ахыска на 23 %, азербайджанцев на 9 %, уйгур на 6,81 %.

Наблюдается уменьшение количества представителей чувашского (38 %), карачаевского (29%), башкирского (26%), балкарского (19,5%) татарского (18%) языков.

Результаты полевых исследований позволили обратить особое внимание на субварианты турецкого языка, представленные на лингвистической карте Казахстана различными вариантами турецкого языка Османского периода, которые вышел из обихода в Турции после реформ Мустафы Кемаль Ататюрка, но сохранились в Казахстане и других Центрально-Азиатских государствах. Турецкое население в 1944 г. была насильственно депортирована из Грузии - Ахалцихского района, находящегося на территории исторической области Месхетии и Аджарии. Лингвистическое документирование Казахстанского варианта турецкого языка является неотложной задачей [Nevskaya, Tazhibayeva 2015a: 294-296].

Эндогенные миноритарные языки – это языки, на которых говорит языковое меньшинство в данной стране, и, которые не являются мажоритарными в другой стране [Trudgill 2003: 43; Сулейменова 2007: 26]. Как правило, они не имеют собственных административных территорий. В Казахстане это – шорский, караимский, крымчакский, крымско-татарский, тофаларский, долганский (см. Таблицу №2) www.stat.gov.kz.

Таблица №2

Количественное соотношение носителей тюркских эндогенных языков

	1970	1979	1989	1999	2009	
Общее кол-во населения	13.026.274	14.709.508	16.222.324	14.981.281	16.009.597	
1. Долгане		17	18	56	25	3
2. Караимы		50	33	33	28	231
3. Крымчаки		42	87	61	20	35
4. Крымские татары	2.023		834	3.125	1.007	1.532
6. Шорцы		215	381	382	212	96

Среди представителей эндогенной группы языков рост на 52 % наблюдается у крымских татар. Рост населения в восемь раз наблюдается у караимов; число крымчаков увеличилось на 75 %. Большое удивление вызывает большое числа караимов, крымчаков. В нашей базе зафиксирован лишь один случай, когда респондент идентифицирует себя караимом по самоопределению, нет данных о крымчаках. В ходе проведения полевых исследований нам не удалось найти представителей этих редких этносов. С лингвистической арены Казахстана уже исчез тофаларский язык, осталось только три представителя из числа долган, на 55% уменьшилось число представителей шорского этноса.

В настоящее время задача исследования истории, языка, литературы и культуры тюркских этносов, проживающих на территории Казахстана, особенно актуальна. Востребованным и необходимым является изучение истории формирования современных тюркских языков, а также отдельных ветвей тюркской семьи языков, исследование их развития в условиях тесных контактов с казахским, русским и другими языками, а также взаимовлияния культур тюркских этносов в Казахстане. Вызывает интерес языковое

строительство тюркских этносов Казахстана, проживающих за пределами исторической родины, и анализ социолингвистических факторов, влияющих на развитие этих языков [см. Тажибаева 2014].

Отдельную перспективную тему для исследования казахстанскими лингвистами представляет собой процесс взаимодействия между близкородственными тюркскими языками и литературами на фоне доминирования одного языка.

Все эти направления исследования являются уже сейчас полем деятельности как казахских, так и зарубежных тюркологов. Научные контакты в области тюркского языкознания между тюркологами Германии, России, Казахстана представляет собой образец сотрудничества научных школ и традиций.

В настоящее время казахстанско-германское сотрудничество укрепляется благодаря совместному международному проекту “Взаимодействие тюркских языков и культур в постсоветском Казахстане”, который выполняется группой ученых Свободного университета г. Берлина (Берлин, ФРГ), Университетом г. Франкфурта (ФРГ), ЕНУ им. Л. Н. Гумилева (Астана, РК), ТарГУ им. М. Х. Дулати (Тараз, РК) при грантовой поддержке фонда “Volkswagen”. В проекте особое внимание исследователей уделяется решению следующих проблем:

- языковое строительство тюркских этносов, проживающих на территории РК;
- социолингвистические факторы, влияющие на развитие этих языков;
- документирование исчезающих тюркских языков [Невская, Тажибаева 2014: 30-31].

Наибольший интерес представляет происходящий в современном Казахстане процесс взаимодействия между близкородственными тюркскими языками [см. Шайердинова, Тажибаева, Невская, 2015]. Этот процесс принципиально схож с аналогичными процессами, происходившими в тюркском мире на протяжении столетий и приведшими в итоге к формированию современных групп тюркских языков, не поддающихся однозначной последовательной классификации ни по одному из базовых признаков, которые в разное время принимались за основу различными специалистами-тюркологами.

Особая значимость проекта заключается в установлении междисциплинарных связей, в привлечении в научное исследование тюркских языков Казахстана студентов, магистрантов, докторантов специальностей тюркология, информационные системы ЕНУ им. Л.Н.Гумилева, а также немецких студентов, магистрантов Франкфуртского университета [см. Nevskaya, Tazhibayeva 2015b]. Это особенно важно для дальнейшего плодотворного сотрудничества, укрепления международных связей и продолжения научных традиций тюркского языкознания и тюркологии в целом.

Одним из важных этапов проекта является проведение социолингвистического исследования. С этой целью была составлена анкета, включающие следующие разделы:

- самоидентификация, этническая и национальная принадлежность по паспорту;
- определение национальности в межнациональных браках; история семьи (история миграции семьи);
- языковая компетентность, использование языков в различных ситуациях;
- поддержка государством языков тюркских этносов (преподавание в школе, средства массовой информации и др.);
- уровень образованности и др. [Nevskaya, Tazhibayeva 2015 b: 324-327]

В рамках проекта создается электронный корпус полученных во время опроса данных, который доступен для всех тюркологов, а также для всех, кто интересуется данной проблематикой, в Интернет ресурсах. Необходимо отметить, что это первая электронная база социолингвистических данных по тюркским этносам.

Анкетный опрос проводился студентами кафедры тюркологии ЕНУ им. Л. Н. Гумилева. Они провели анкетирование более 2000 респондентов из разных регионов Казахстана. Разработан сайт для анкетирования респондентов, где можно самостоятельно заполнить анкету. Все ответы, как полученные путём непосредственного общения с респондентами, так и путём самостоятельного заполнения анкеты представителями тюркских этносов Казахстана, вносятся в электронную базу данных. Результаты анкетирования опубликованы на сайте в виде графиков и диаграмм с применением фильтров.

Создание базы данных представляет планомерную работу, состоящую из четырех этапов: сбора данных, хранения информации, обработки информации, визуализации данных [Fedotov, Tussupov, 2015; Tussupov, Sambetpayeva, 2015].

1. Сбор данных с помощью анкетирования респондентов.

Анкеты заполнялись в бумажной форме и на сайте, который был специально разработан для проведения анкетирования. Респонденты заполняли анкеты на бумажных носителях в связи с отсутствием доступа к интернетсети в сельской местности, где проживает большинство представителей тюркских этносов Казахстана, либо это были представители старшего поколения (ветераны, пенсионеры), часто не владеющие компьютером. Ответы респондентов вносились в электронную базу данных, которая хранится на сервере Google Drive.

Google Drive – это облачное хранение данных, в состав Google Диска, в которые входят Google Документы, Таблицы и Презентации – набор офисных приложений для совместной работы над текстовыми документами, электронными таблицами, презентациями, чертежами, веб-формами и другими файлами.

Сайт для анкетирования был разработан с использованием веб-форм студентами и магистрантами кафедры информационных систем ЕНУ им. Л.Н. Гумилева.

Веб-форма на сайте представляет собой аналог бумажной формы, анкеты, бланка и опросного листа. В формах имеются поля, предназначенные для заполнения, а также списки и переключатели, которые позволяют пользователю выбрать один или несколько элементов. Форма может осуществлять свою работу только в том случае, если она поддерживается веб-сервером. Использование веб-форм, как правило, предусматривает использование на сайте баз данных.

Анкеты можно было также заполнить на сайте www.tuyrki.weebly.com.

Из числа опрошенных лиц, участвовавших в социолингвистическом исследовании, 54% составили женщины и 45% мужчины. Такое процентное соотношение среди респондентов свидетельствует о том, что женская часть населения легче вступает в контакт и готова представить информацию о себе и о своем этносе.

В анкетировании участвовали респонденты в возрасте от 10 до 98 лет. Для полноты охвата и репрезентативности информации к анкетированию были привлечены лица почти каждой возрастной категории. Наибольшая доля опрошенных респондентов приходится на группы в возрасте 20-75 лет. По остальным возрастным группам численность опрошенных респондентов значительно ниже. Самому старшему респонденту 98 лет, самому младшему 10 лет. Возрастной разрыв опрошенных в исследуемой совокупности составил 88 лет. Это позволяет провести исследование изменения социокультурного взаимодействия тюркских языков и культур почти в вековом интервале исторического развития.

Результаты социолингвистического исследования обрабатываются и синхронизируются в базе данных, доступной для широкой аудитории в интернет-ресурсах.

В качестве базы данных для международного проекта было использовано приложение Google Таблицы, которые хранятся на облачном сервере Google Drive [<http://www.uznaem-kak.ru/baza-dannyx>].

2. Хранение результатов опроса.

После завершения анкетирования, ответ респондента записывается в базу данных, представленной в табличной форме. Все данные хранятся на сервере Google Drive.

На рисунке 1 представлена табличная форма хранения ответов.

Отметка времени	Жасыңыз (Сіздің жан)	Қай елдің азаматы бол	Сіз қай этносқа жата	Төлсұжат бойынша / По паспорту	Өзіндік тағым бойы	Ұлты / Национальн	Сіздің жеңілдігі / Би
17.09.2013 15:20:37	46	Қазақстан	түрки	түрки хемшілші	түрки	түрки	түрки хемшілші
17.09.2013 15:59:51	46	Қазақстан	қара-леңғи	қара-леңғи	қара-леңғи	қара-леңғи	қара-леңғи
17.09.2013 17:07:58	19	Қазақстан	қара-леңғи	қара-леңғи	қара-леңғи	қара-леңғи	қара-леңғи
17.09.2013 17:32:55	75	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 10:22:22	19	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 11:11:31	52	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 11:34:32	48	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 11:47:17	73	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 11:59:08	75	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 12:28:44	33	Қазақстан	узбеки	узбеки	узбеки	узбеки	узбеки
18.09.2013 12:40:06	26	Қазақстан	узбеки	узбеки	узбеки	узбеки	узбеки
18.09.2013 12:48:56	27	Қазақстан	узбеки	узбеки	узбеки	узбеки	узбеки
18.09.2013 13:24:51	28	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 15:24:18	21	Қазақстан	узбеки	узбеки	узбеки	узбеки	узбеки
18.09.2013 16:05:26	14	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 18:20:45	55	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 18:41:21	75	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 18:57:32	25	Қазақстан	татары	татары	татары	татары	татары
18.09.2013 19:11:14	40	Қазақстан	татары	татары	татары	татары	татары

Рисунок 1. Табличная форма хранения ответов

3. Обработка данных.

В базе данных для поиска ответов на запросы исследователя можно осуществить фильтрацию для быстрого и простого способа нахождения подмножества данных и работы с ним в диапазоне ячеек или в столбце таблицы. В отфильтрованных данных отображаются только строки, соответствующие заданным условиям, а ненужные строки скрываются. Фильтрация позволяет отфильтровать несколько столбцов. Можно применять одновременно несколько фильтров, при этом каждый следующий фильтр добавляется к результатам и дополнительно сужает подмножество данных. В списке, содержащем текстовые значения, можно установить или снять флажки текстовых значений в зависимости от условий, по которым нужно выполнить фильтр [<https://support.office.com/ru-ru/article>].

4. Визуализация данных.

Табличная форма для отображения результатов с фильтрами является недостаточно наглядной, поэтому необходима визуализация данных в виде диаграмм и графиков с применением того же фильтра. Для этих целей была разработана визуализация данных при помощи инструментов Google Chart API, основанных на языке программирования Java Script.

Google Chart API — это бесплатный сервис от компании Google, который позволяет разработчикам веб-приложения создавать изображения диаграмм путем создания элемента `` с атрибутом `src`, указывающего на ссылку URL, которая включает в себя данные диаграммы, ярлыки и другую информацию для строки запроса. Google Chart API предлагает быстрый и легкий способ создания множества различных видов диаграмм для веб-сайта, а также создавать диаграммы на основе информации из базы данных [<http://www.interface.ru/home.asp?artId=22195>].

Основной ссылкой на Google Chart API является <http://chart.apis.google.com/chart?>. Параметры, которые определяют отображение диаграммы, следуют после символа `?`. Существует множество параметров, которые можно указать посредством строки запроса. Единственными обязательными параметрами являются размер диаграммы (`chs`), данные диаграммы (`chd`) и тип диаграммы (`cht`)

На рисунке 2 представлена «оболочка» сайта, в котором выполнена визуализация данных с использованием инструментов Google Chart API и языка программирования Java Script.



Рисунок 2 Сайт с использованием инструментов Google Chart API и языка программирования Java Script

Все данные были разделены на 6 подгрупп для удобства проведения визуализации. Если выбирается подгруппа национальность, то можно применить фильтр, который автоматически выводит диаграммы.

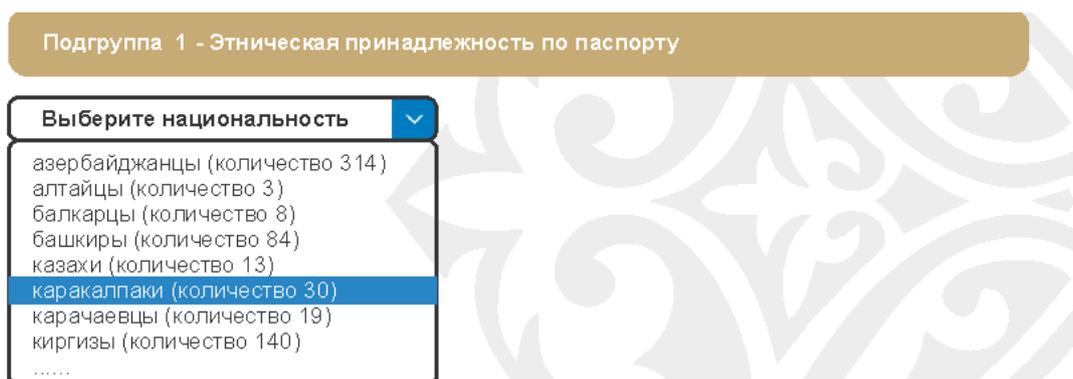


Рисунок 3. Визуализация подгруппы 1 – Этническая принадлежность по паспорту.

В качестве инструмента для построения диаграмм был выбран Google Chart API, позволяющий создавать диаграммы на основе информации из базы данных. На рисунке 4 представлены диаграммы, в которых отражены вопросы, касающиеся истории семьи, местожительства в Казахстане, религиозной принадлежности, занимаемой должности и др. В диаграммах

отражены ответы, которые были даны респондентами, например, указаны только восемь регионов из четырнадцати областей Казахстана; представлены шесть наиболее часто отмеченных названий профессий, остальные вошли в группу «другие». При наведении мышкой на интересующую часть диаграммы отображается процентное соотношение данных.

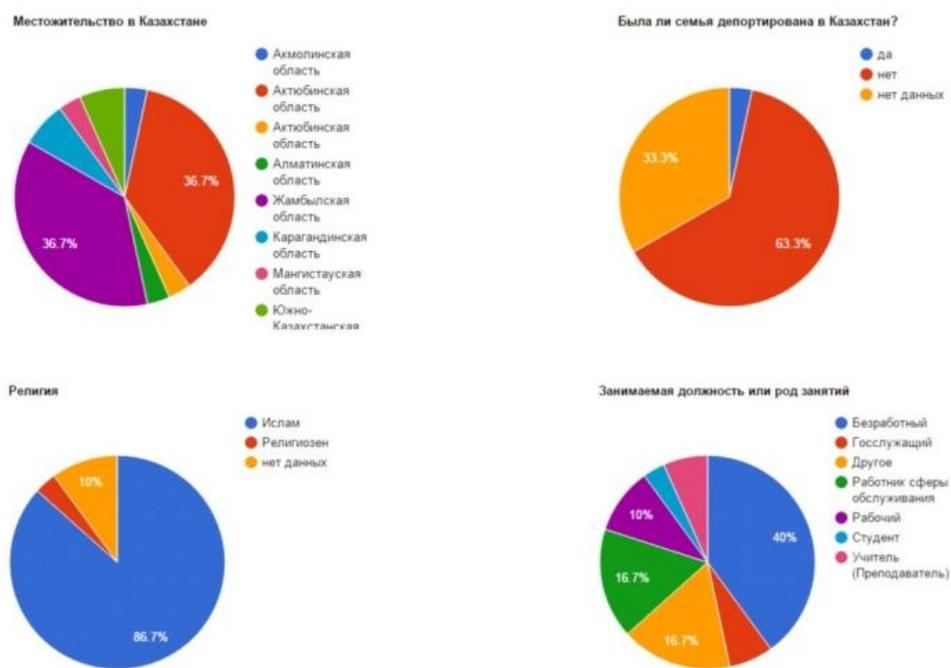


Рисунок 4. Визуализация диаграмм по заданному вопросу.

Таким образом, тюркский мир Казахстана представляет большой научный интерес для исследователей-тюркологов. Создание базы данных позволяют эффективно использовать лингвистический и этнографический материал. Данные социолингвистического опроса намечают новые перспективы тюркологических и социолингвистических исследований в РК.

Список литературы и электронных ресурсов:

1. Сулейменова Э.Д. Шаймерденова Н.Ж. Аханова Д.Х. Языки народов Казахстана. Социолингвистический справочник. – Астана: Изд-во «Арман-ПВ», 2007. – 304 с.
2. Trudgill P. Glossary of Social Linguistics. – Edinberg University Press, 2003.
3. Итоги национальной переписи населения 2009 года в Республике Казахстан Статистический сборник /Под ред. А.Смаилова/ Астана, 2010. www.stat.gov.kz)
4. Nevskaya I., Tazhibayeva S. 2015a. Turkic Languages of Kazakhstan: Problems and research perspectives. In: Lars Johanson (ed.) Turkic Languages.

Volume 16, 2014. Numbers 1/2. – P. 289-302. – Harrassowitz Verlag. Weisbaden, 2015.

5. *Тажыбаева С.Ж.* Қазақстанда тұратын түркі этностарының тілдерін зерттеудің келелі мәселері - Л.Н.Гумилев атындағы ЕҰУ Хабаршы ғылыми журналы – Астана, 2014 - №3 (100) 2014-10-31 I бөлім - С. 118-124.

6. *Невская И.А., Тажибаева С.Ж.* Современная тюркология и международное сотрудничество на евразийском пространстве. - Global-Turk. International Journal for Information and Analytics. – Turkic Academy. - №1, 2014. - . С. 19-34

7. *Шаймердинова Н.Г., Тажибаева С.Ж., Невская И.А. Schönig С.* Татарская диаспора Казахстана: первые результаты социокультурного анализа. - *Қазақ лексикографиясы: тарихы, тәжірибесі, болашағы.* Халықаралық ғылыми теориялық конференция материалдары – Алматы, 2015. – С. 199-203.

8. *Nevskaaya I., Tazhibayeva S.* 2015 b. Sociolinguistic situation of Oguz Turks in post-Soviet Kazakhstan In: *Oguzlar. Dilleri, Tarihleri ve Kulturleri* - Ankara, 2015 – P.321-334. - ISSN 978-975-491-405-4.

9. *Fedotov, J. Tussupov, M. Sambetbayeva, I. Idrisova, A. Yerimbetova.* 2015. Development and Implementation of a Morphological Model of Kazakh Language. In: *Eurasian Journal OF Mathematical AND Computer Applications* (ISSN 2306–6172). Volume 3, Issue 3 (2015), 69– 79; *Tussupov, J., Sametbayeva, M., Jetpisov, K., Mukhanova, A.*, 2015. Formalization of Ontology and its Algorithmic Properties. In: *15th Congress of Logic, Methodology and Philosophy of Science (LC-2015).* Book of abstracts. – University of Helsinki, P. 746. (ISSN 1079-8986)

10. www.stat.gov.kz

11. <http://www.eng.stat.kz>

12. <http://www.uznaem-kak.ru/baza-dannyx>

13. <https://support.office.com/ru-ru/article>

14. <http://www.interface.ru/home.asp?artId=22195>

15. <http://chart.apis.google.com/chart?>

Түйіндеме

Мақала қазіргі кезеңдегі Қазақстандағы түркі тілдерінің әлеуметтік лингвистикалық зерттеулері мен оның базасын жасаудың алғашқы тәжірибелеріне арналған

Resume

The article is devoted to sociolinguistic research on Turkic languages in contemporary Kazakhstan and the first experience of data base creation.