

## Структура и единицы морфоанализатора узбекского языка

Б.Р. Менглиев, Ш.М. Хамроева

Ташкентский государственный университет узбекского языка и литературы  
им. Алишера Навои

bakhtiyormengliiev@gmail.com, hamroyeva81@mail.ru

### Аннотация

В статье рассматриваются вопросы создания морфологического анализатора узбекского языка, в том числе, единицы которые составляет морфоанализатора. Для достижения развития узбекского языка в современных информационных технологиях, необходимо аргументировать изучение таких современных вопросов, как языковые корпуса, электронные переводчики, тезаурусы, орфографические корректоры, инструменты автоматической обработки (лингвистические анализаторы) и их лингвистическое обеспечение на основе современных научных принципов. В статье рассматриваются вышеуказанные вопросы и решения вопросов создания морфологического анализатора узбекского языка, в том числе, структура и состав морфоанализатора узбекского языка.

**Ключевые слова:** морфологический анализатор, узбекский язык, граммема, дериватема, квазиграммема

**Библиографическая ссылка:** Менглиев Б.Р., Хамроева Ш.М. Единицы морфоанализатора узбекского языка // Компьютерная лингвистика и вычислительные онтологии. Выпуск 5 (Труды XXIV Международной объединенной научной конференции «Интернет и современное общество», IMS-2021, Санкт-Петербург, 24 – 26 июня 2021 г. Сборник научных статей). — СПб.: Университет ИТМО, 2021. С. 51-60. DOI: 10.17586/2541-9781-2021-5-51-60

### Введение

Одним из первых этапов компьютерной обработки естественного языка является выполнение морфологического анализа, а модуль морфоанализа является основным компонентом большинства программ анализа текста. В мировой компьютерной лингвистике автоматический морфоанализ изучается с 1950-х годов [11], в том числе применительно к узбекскому языку. Создание морфоанализатора узбекского языка начинается с анализа его единиц, состава и структуры. Чтобы запустить морфологического анализатора (МА), обычно формируются определенные компоненты. Составной частью МА является справочные базы данных: база данных аффиксальных морфем; база основных морфем; база данных флекций и др. Также разрабатывается различные лингвистические правила: правила организации синтаксической структуры, правила совместимости алломорфов [2; стр. 73], правила разрешения неоднозначности и др. Н.А. Исраилова, П.С. Бакасова описывают структуру морфологической базы данных следующим образом:

- 1) морфологическая база данных должна содержать всю информацию, необходимую для процесса морфологического анализа и синтеза;
- 2) существующие флективные случаи в языке, а также возникающие в результате фонетические изменения корня и дополнительных границ должны быть отражены в базе данных;

- 3) морфологическая база данных должна содержать информацию о постоянных, супплетивных лексемах, а также об обычных синтаксических формах;
- 4) в базе данных должна быть база омонимических лексем, а также омонимов, образованных под влиянием полной и грамматической формы [16; 110-117].

Исходя из приведенных выше соображений, алгоритм работы морфологического анализатора можно представить в виде следующей схемы (рис. 1).



Рис. 1. Состав и структура морфологического анализатора

Эти составные части могут сильно варьироваться в зависимости от языка и от конкретной реализации. Кроме того, в последнее время активно разрабатываются морфоанализаторы, базирующиеся на методах машинного обучения и на нейронных сетях.

Исходя из вышеуказанных основные компоненты морфологического анализатора узбекского языка можно указать следующим образом:

- 1) словарь языка (словарь, содержащий частеречные пометы и чередования основ, не описанные фонологическими правилами);
- 2) компьютерная модель словоформы, опирающаяся на адекватное грамматическое описание (ориентированное на автоматический анализ);
- 3) набор правил сочетаемости, включающий правила сочетаемости аффиксов в пределах словоформы и фонетические правила выбора алломорфов конкретного аффикса.

## 2. Основы разработки морфологического анализатора узбекского языка

В данной статье мы рассмотрим лингвистическое обеспечение морфоанализатора узбекского языка. Узбекский язык носит агглютинативный характер, возможность выражения грамматического значения с помощью аффиксов широка. Поэтому унификация грамматических форм важна при построении морфологического анализатора.

Агглютинативная словоформа образуется путём присоединения к основе в строгом порядке однозначных стандартных аффиксов; границы морфем отчетливы, фонетические изменения на стыках морфем подчиняются строгим правилам. Но попытки построить парадигму конкретного слова демонстрируют ее чрезвычайную сложность и многоместность, что обусловлено большим числом словоизменятельных аффиксов. Это подталкивает нас к построению морфологического анализатора, учитывающего все возможные в тюркских языках комбинации морфем. Существующие тюркские парсеры строят редуцированные парадигмы, ограничивая число словоформ примерно до трехсот, что пагубно сказывается в качестве их работы. Ниже описываем свойства агглютинативных языков.

Морфологические признаки агглютинации:

1. Корень слова – в именительном падеже выступает в чистом виде, таким образом, является центром всей парадигмы склонения.
2. Между морфемами четко сохраняется граница.
3. Строгая последовательность присоединения аффиксов.

Фонетические признаки агглютинации:

1. Наличие сингармонизма.
2. Фиксированное ударение, которое способствует сохранению фонетической целостности слова.

Синтаксические признаки агглютинации:

1. Твёрдый порядок слов в предложении.
2. Определение находится перед определяемым словом.
3. Дополняющее слово находится перед дополняемым словом.
4. Сказуемое в конце предложения [8: 103-105].

Разработка морфологических анализаторов тюркских языков началась в 60-х годах прошлого века. Главная особенность первых анализаторов заключалась в том, то что они не были рассчитаны только на один язык: параллельно морфоанализатор можно было адаптировать к другому языку [11; 98-105]. Если система правил (алгоритм) базируется на лексических и аффиксальных словарях, представляется возможным использовать ее как анализатор другого языка, поменяв эти словари; нет необходимости писать отдельный программный код для создания анализатора другого языка. Однако все оказалось не так просто. С тех пор изменились технологии, созданы морфологические анализаторы отдельных тюркских языков, увеличился объем словарей, скорость обработки данных и снова предпринимаются попытки создания универсальных морфологических анализаторов. Хотя с начала разработки морфоанализаторов тюркских языков прошло более 50 лет, во всех тюркских языках эта область все еще остается актуальной и находится на разном уровне. Подтверждением этому может служить обзор работ над татарским [3; 10], башкирским [9], казахским [12], чувашским [6], турецким [1], хакасским [5] языками, а также над идеей создания универсального морфологического анализатора [4].

В литературе представлены такие подходы к автоматическому морфологическому анализу, как стемминг (1), лексический анализ словоформы на основе словаря (2), бессловарный анализ [11; 65].

### 3. Морфологическая специализация для узбекского языка

Любой формат имеет структуру. Здесь мы описываем наполнение формата морфологической разметки узбекского языка. Мы уже упоминали о составной части морфоанализатора и напомнили, что словарь баз входит в данный состав. Обсудим этот вопрос ниже.

Также специалисты включают лексикон основных лексем в число компонентов, составляющих работу морфоанализатора [11; 56]. Лексикон основных лексем формируется на основе современного литературного языка. Обычно словари, охватывающие относительно полный набор языковых словарей, составляют основу

лексикона. Наши наблюдения основаны на требованиях инструментария HFST лексикона традиционных морфоанализаторов указывают на то, что он состоит из 9 следующих групп: Nouns (существительные), Verbs (глаголы), Adjectives (прилагательные), Adverbs (наречие), Pronouns (Местоимения), Numerals (Числительные), Postpositions (вспомогательные послелог), Conjunctions (союзы), Exclamations (Междометия).

Прежде всего опишем структуру частей речи узбекского языка (Табл. 1).

**Таблица 1.** Таблица тегов частей речи узбекского языка

№	Категория	Код
1	Глагол	V
2	Существительное	N
3	Прилагательное	ADJ
4	Числительное	NUM
5	Наречие	ADV
6	Местоимение	SPRO
7	Имитирующие слова	IMIT
8	Модальные слова	MOD
9	Междометие	INTJ
10	Союз	conjunction
11	Послелог	POST
12	Частица	PART

«Толковый словарь узбекского языка» является важным ресурсом для определения того, к каким частям речи относятся лексемы. Комбинации аффиксов описаны в «Морфологическом словаре узбекского языка» [27]. Для работы любого анализатора требуется базовый словарь – лексикон. Для морфоанализатора узбекского языка такой машинный словарь может быть создан на базе «Узбекского орфографического словаря» [13] или пятитомного «Толкового словаря узбекского языка» [14, 15], который содержит 85000 слов.

Ниже приведем для каждой части речи ее атрибуты (категории) и их значения (граммемы). Атрибуты существительных узбекского языка и их значения (граммемы) приведем ниже (см. Табл. 2).

**Таблица 2.** Категории и значения существительных узбекского языка

№	Грамматическая категория	Грамматическое значение	тег
1	Имена собственные		NOUN1
2	Имена предметные		NOUN2
3	Абстрактное существительное		DERIV.ABSTR
4	Категория падежа	основной падеж родительный падеж дательный падеж винительный падеж исходный падеж местно-временной падеж	NOM GEN DAT ACC ABL LOC
5	Категория лица	первое лицо, ед. число второе лицо, ед. число третье лицо, ед. число первое лицо, множ. число второе лицо, множ. число третье лицо, множ. число	NOM GEN DAT ACC ABL LOC

Атрибуты глаголов узбекского языка и их значения (граммемы) приведем ниже (см. Табл. 3).

**Таблица 3.** Категории и значения глагола

№	Грамматическая категория	Грамматическое значение	тег
1	Глагол Типы глаголов по лексико-грамматическим признакам	независимый глагол зависимый глагол • вспомогательный глагол • неполный глагол	V V-INDEP V-DEP V-DEP1 V-DEP2
2	Лексико-семантические группы	глагол действия речевой глагол глагол состояний глагол эффективной деятельности созерцательный глагол глагол отношения	actionV speechV stateV effectiveV contempV relationshipV
3	Типы глаголов по отношению действия к объекту	Переходный глагол непереходный глагол	TransitiveV intransitiveV
4	(аспекты глагола) Реальные/ирреальные	Положительная форма отрицательная форма	POZ NEG
5	Категория залога	активный залог Возвратный залог пассивный залог Понудительный залог Взаимно-совместный залог	ACT REFL PASS CAUS RECP
6	финитность	Имя действия причастие деепричастие	SUP ParI GER
7	Категория наклонения	Изъявительное наклонение Условное наклонение Повелительное наклонение Желательное наклонение	ACTv COND IMP OPT
8	Категория времени	Прошедшее время Настоящее время Будущее время	PST PRES FUT
9	Категория число и лицо	первое лицо, ед. число второе лицо, ед. число третье лицо, ед. число первое лицо, множ. число второе лицо, множ. число третье лицо, множ. число	POSS 1SG POSS 2SG POSS 3SG POSS 1PL POSS 2PL POSS 3PL

Атрибуты других частей речи узбекского языка и их значения (граммемы) приведем ниже (см. Табл. 4, 5, 6, 7).

**Таблица 4.** Категории и значения прилагательных

№	Грамматическая категория	Грамматическое значение	код
1	Имена прилагательные непроизводные	Non-derivative adjectives	DER.ADJ
2	имена прилагательные производные	derivative adjectives	N.DER.ADJ
3	Степени сравнения	Простая тепень Сравнительная степень (comparative) Превосходная степень	COMP

**Таблица 5.** Категории и значения местоимение

№	Грамматическая категория	Код
1	Местоимение	SPRO
2	Личные местоимение	SPRO1
3	Указательные местоимение	SPRO2
4	Возвратные местоимение	SPRO3
5	вопросительные местоимение	SPRO4
6	притяжательные самостоятельные местоимение	SPRO5
7	Определительные местоимение	SPRO6
8	неопределенные местоимение	SPRO7

**Таблица 6.** Категории и значения числительных

№	Грамматическая категория	Грамматическое значение	Код
1	числительное	количественные числительные	DIST
		штучные числительные	PIS.N
		порядковые числительные	ORD
		разделительные числительные	DIVIS
		дробные числительные	DIVIS
		собирательные числительные	COL
		неопределенно-количественные числительные	APRX
		нумеративы	NUMer

**Таблица 7.** Категории и код вспомогательных слов

№	Грамматическая категория	Код
1	Частица (particle)	PART
2	Послелог (postposition)	POST
3	Союз (conjunction)	CONJ
4	Междометие (interjunction)	INTJ
5	Имитирующие слова (imitation)	IMIT
6	модальные слова (modal words)	MOD

### 3. Особенности морфологического анализа узбекского языка

Моделирование и аранжировка грамматических форм узбекского языка. Порядок грамматической форм в узбекском слове имеет определенную закономерность: корень + словообразовательный суффикс + лексический суффикс + синтаксический суффикс. Словесная среда грамматических форм определяет его значение/функцию. Из-за наличия грамматических форм с разными формами и средами слов, среды словообразования и словоизменения уникальны с точки зрения глагола, существительного, прилагательного, числительного, местоимения, наречия, поэтому их словесная среда должна быть специально смоделирована для морфоанализатора. Автоматический морфоанализ не может быть достигнут путем моделирования грамматических форм, которые имеют одинаковую форму и словесную среду: они должны быть семантически/синтаксически проанализированы при автоматической обработке.

Анализ и моделирование круга лексических формообразующих морфем в словах обеспечивает его отличие от словообразовательных и словоизменительных. Словоизменительные морфемы, составляющие существительные и глаголы, образуют грамматические омонимы.

Для того, чтобы морфоанализатор мог идентифицировать производное слово, образующую часть и образующую основу, необходимо иметь точную информацию о структуре производного слова.

В словаре основ анализатора все основы узбекского языка хранятся с информацией о категории: к какой части речи они относятся. На основе моделей словообразования анализатор определяет три вещи: основу слова, основу образования и словообразовательный аффикс.

Хотя узбекский язык носит агглютинативный характер и подчиняется строгим правилам, в его грамматических правилах есть некоторые исключения. Например, уровень превосходной степени прилагательных появляется от первого слога и записывается с дефисом перед основанием: кип-қизил, қоп-қора, яп-янги и т.д. Для формирования лингвистического обеспечения морфоанализатора необходимо выявить такие формы в узбекском языке, сформировать список, включить в базу прилагательных лексикона как словарь исключений.

Некоторые вопросы разрешение неоднозначности грамматических форм. Среди единиц тюркского языка в связи с наличием единиц со сложной структурой слова, при описании грамматических знаков наблюдаются случаи неправильной (автоматической) разметки (тегирования). Примером является морфема -ган, которую можно интерпретировать как две разные морфемы: одна используется для обозначения категории прошедшего времени, а другая для обозначения категории причастия. По результатам анализа морфоанализатор предлагает два варианта. Чтобы найти ответы на данные вопросы, мы сосредоточимся на опыте использования различных типов классификации языковых единиц для решения проблемы.

Классификация морфем на основе параметров сложения и вычитания. В тюркских языках аффиксальные морфемы, относящиеся к одной грамматической категории, могут быть представлены как синтаксическими, так и аналитическими единицами, то есть в одном тюркском языке они пишутся слитно, а в другом – раздельно. Например, в татарском языке грамматическая категория белән, которая участвует в соединении слова перед собой со следующим словом, представлена только вспомогательным предлогом и записывается отдельно от слова перед собой [2]. В узбекском языке (а также в казахском и турецком языках) данная грамматическая категория, то есть значение средства выражения, выражается как падежом, так и вспомогательным предлогом. Сравните:

– по-казахски: аффикс *-бен* с *-бен*, *-мен*, *-пен* алломорфами и с вспомогательные средствами *бенен*, *менен*, *пенен*;

– по-турецки: с аффиксом *-la* и с вспомогательным средством *ile*.

– по-узбекски: с аффиксом *-да* и с вспомогательные средствами *билан*, *ила* (с алломорфом *-ла*).

Обратим внимание на значенииЕ вспомогательных средств (см. примеры из узбекского толкового словаря).

БИЛАН вспомогательный предлог. 1 Используется с разными словами в именительном падеже: 1) представляет собой совместное действие. *Оттам билан келдим. Мт Кўчаларни, қишлоқни, айниқса, Маҳваш билан бирга дарс тайёрлаган жойларни... жуда-жуда қўмсадим.* Т.Ашуров, Оқ от; 2) относится к тому, что является средством для действия или события. *Ошининг таъми туз билан, одам таъми сўз билан. Пословица. Сут билан кирган – жон билан чиқар. Пословица. т т Косада лагмон кўтариб кирган Марғуба бир қўли билан шишани ушлади.* С.Зуннунова, Кўк чироқлар. *Инсон хатони кўпроқ ўз тили билан содир қилади. Из газеты [13;257].*

Как видно из примеров, вторым значением вспомогательного выражения *билан* является «означает что-то, что является средством к действию или событию», то что значение средств выражается вспомогательным предлогом.

ИЛА род. падеж. билан. *Бой ила хизматчи. тт [Кумуш] ..мактубни ўқиб тамом қилди-да, жонсиз бир товуш ила «куятсиз» деб қичқирди ва ерга йиқилиб, ўздан кетди.* А. Кадири, Ўтган кунлар [14; 187-188].

В данном примере предлог *ила* также представлен вспомогательным средством.

Вспомогательные слова, выражающие грамматическое значение, представлены как суффиксом, так и словом. Следовательно, в анализаторе различаются такие показатели как лексическая единица (L); аналитическая форма (AF) и аналитическая конструкция (AK) [7; 76].

## Выводы

При разработке морфоанализатора формируются такие компоненты как, база данных аффиксальных морфем, база основных морфем; правила классификации; правила совместимости алломорфов. В морфоанализаторе база данных имеет большое значение для охвата всей информации, необходимой для процесса анализа/синтеза; случаи существующей флекции в языке – отражение фонетических изменений границы корня и суффикса; охват информации о дополнительных, инвариантных лексемах, наряду с традиционными словоизменителями; также требования, как наличие базы из омонимии лексем (а также суффиксов, образованных под влиянием полных и грамматических форм).

Мы разработали морфологическую спецификацию для узбекского языка. Она содержит перечень морфологических категорий для всех частей речи. При этом часть вопросов остались нерешенными. Их разрешение требует статистического анализа языкового материала и будет выполнено на стадии пилотного проекта.

## Литература

- [1] Kemal Oflazer. Two-level Description of Turkish Morphology // *Literary and Linguistic Computing*, 1994. Vol. 9. № 2.
- [2] Архангельский Т.А. Принципы построения морфологического парсера для разноструктурных языков: диссертация кандидата филологических наук. Москва, 2012.
- [3] Гатиатуллин А.Р., Баширов А.М. Морфологический анализатор тюркских словоформ на базе структурно-функциональной модели тюркской морфемы // Пятая Международная конференция по компьютерной обработке тюркских языков «TurkLang 2017». Труды конференции. В 2-х томах. Т 2. Казань: Издательство Академии наук Республики Татарстан, 2017. С. 50-71.
- [4] Дрейзин Ф.А. Об алгоритмизации составление алгоритма анализа языка (на примере морфологии агглютинативного узбекского языка) // *Научные труды Ташкентского университета*. 1961. вып. 189. Матем. науки. С. 121.
- [5] Дыбо А.В., Шеймович А.В. Автоматический морфологический анализ для корпусов тюркских языков // *Филология и культура*. 2014. № 2.
- [6] Желтов П.В. Морфологический анализатор чувашского языка // *Материалы Международной конференции студентов и аспирантов по фундаментальным наукам «Ломоносов 2002»*. М., 2002.
- [7] Марчук Ю. Компьютерная лингвистика. Москва, 2006. С. 65.
- [8] Плунгян В.А. Классификация морфологических значений // *Общая морфология: Введение в проблематику: Учебное пособие*. – Изд. 2-е, исправленное. – М.: Едиториал УРСС, 2003. С. 107. 384 с.
- [9] Сиразитдинов З.А. Алгоритмическая грамматика словоизменения башкирского языка // <http://mfbl.ru/bashdb/algram/algram.htm>.
- [10] Сулейманов Д.Ш., Гильмуллин А.А., Гильмуллин Р.А. База морфотактических правил для татарского глагола как основа двухуровневого морфологического анализатора // *Сборник трудов Международного семинара «Диалог»*. Казань, 1998. С. 597-609.



- [11] Формальные модели и программные инструменты компьютерной обработки татарского языка / Р.Р.Гатауллин, А.Р.Гатиатуллин, О.А.Невროзова, Д.Р.Мухамедшин, Д.Ш.Сулейманов, Б.Э.Хакимов, А.Ф.Хусаинов. Академия наук РТ, Институт прикладной семиотики АН РТ. Казань: Изд-во Академии наук Рт, 2019. С.39-40.
- [12] Шарипбаев А.А., Бекманова Г., Ергеш Б.Ж., Бурибаева А.К., Карабалаева М.Х. Интеллектуальный морфологический анализатор, основанный на семантических сетях // Материалы международной научно-технической конференции «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS-2012). Минск, БГУИР, 16-18 февраля 2012 г. С. 397–400.
- [13] Ўзбек тилининг имло луғати. (Тузувчилар: Э.А.Бегматов, А.П.Мадвалиев; Н.Махмудов тахрири остида). Тошкент: Akademnashr, 2013. 520 б.
- [14] Ўзбек тилининг изоҳли луғати. I жилд. Тошкент: ЎзМЭ, 2006. Б.257.
- [15] Ўзбек тилининг изоҳли луғати. II жилд. Тошкент: ЎзМЭ, 2006. Б.187-188.

## Structure and Units of the Uzbek Language Morphoanalyzer

B.R. Mengliev, Sh.M. Khamraeva

Tashkent State University of Uzbek Language and Literature named after Alisher Navoi

The article deals with the issues of creating a morphological analyzer of the Uzbek language, including the units that make up the morphological analyzer. To achieve the Uzbek language in modern information technology, it is necessary to substantiate the study of such modern issues as language corpuses, electronic translators, thesaurus, spelling correctors, the first automatic processing tools (linguistic analyzers) and linguistic support based on modern scientific principles. The article discusses the above issues and the solution of the issues of creating a morphological analyzer of the Uzbek language, including the structure and composition of the morphological analyzer of the Uzbek language.

**Keywords:** morphological analyzer, Uzbek language, grammeme, derivative, quasigrammeme

**Reference for citation:** Mengliev B.R., Khamraeva Sh.M. Structure and Units of the Uzbek Language Morphoanalyzer // Computer Linguistics and Computing Ontologies. Vol. 5 (Proceedings of the XXIV International Joint Scientific Conference «Internet and Modern Society», IMS-2021, St. Petersburg, June 24-26, 2021). - St. Petersburg: ITMO University, 2021. P. 51 – 60. DOI: 10.17586/2541-9781-2021-5-51-60

## Reference

- [1] Kemal Oflazer. Two-level Description of Turkish Morphology // Literary and Linguistic Computing, 1994. Vol. 9. № 2.
- [2] Arhangel'skij T.A. Principy postroeniya morfologicheskogo parsera dlya raznostrukturnykh yazykov: dissertatsiya kandidata filologicheskikh nauk. Moskva, 2012.
- [3] Gatiatullin A.R., Bashirov A.M. Morfologicheskij analizator tyurkskih slovoform na baze strukturno-funkcional'noj modeli tyurkskoj morfemy // Pyataya Mezhdunarodnaya konferenciya po komp'yuternoj obrabotke tyurkskih yazykov «TurkLang 2017». Trudy konferencii. V 2-h tomah. T 2. Kazan': Izdatel'stvo Akademiinauk Respubliki Tatarstan, 2017. S. 50-71.
- [4] Drejzin F.A. Ob algoritimizacii sostavlenie algoritma analiza yazyka (na primere morfologii agglyutinativnogo uzbekskogo yazyka) // Nauchnye trudi Tashkentskogo universiteta. 1961, vyp. 189. Matem. Nauki. S. 121.

- [5] Dybo A.V., Shejmovich A.V. Avtomaticheskij morfoloģicheskiĭ analiz dlya korpusov tyurkskih yazykov // Filologiya i kul'tura. 2014. № 2.
- [6] Zheltov P.V. Morfoloģicheskiĭ analizator chuvashskogo yazyka // Materialy Mezhdunarodnoj konferencii studentov i aspirantov po fundamental'nyĭ naukam «Lomonosov 2002». M., 2002.
- [7] Marchuk YU. Komp'yuternaya lingvistika. Moskva, 2006. S. 65.
- [8] Plungyan V.A. Klassifikaciya morfoloģicheskih znachenij // Obshchaya morfoloģiya: Vvedenie v problematiku: Uchebnoe posobie. – Izd. 2-e, ispravlennoe. – M.: Editorial URSS, 2003. S. 107. 384 s.
- [9] Sirazitdinov Z.A. Algoritmicheskaya grammatika slovoizmeneniya bashkirskogo yazyka // <http://mfbl.ru/bashdb/algram/algram.htm>.
- [10] Sulejmanov D.Sh., Gil'mullin A.A., Gil'mullin R.A. Baza morfotakticheskikh pravil dlya tatarskogo glagola kak osnova dvuhurovnevnogo morfoloģicheskogo analizatora // Sbornik trudov Mezhdunarodnogo seminaru «Dialog». Kazan', 1998. S. 597-609.
- [11] Formal'nye modeli i programnye instrumenty komp'yuternoj obrabotki tatarskogo yazyka / R.R.Gataullin, A.R.Gatiatullin, O.A.Nevrozova, D.R.Muhamedshin, D.Sh.Sulejmanov, B.E.Hakimov, A.F.Husainov. Akademiya nauk RT, Institut prikladnoj semiotiki AN RT. Kazan': Izd-vo Akademii nauk Rt, 2019. S.39-40.
- [12] SHaripbaev A.A., Bekmanova G., Ergesh B.Zh., Buribaeva A.K., Karabalaeva M.H. Intellektual'nyĭ morfoloģicheskiĭ analizator, osnovannyĭ na semanticheskikh setyah // Materialy mezhdunarodnoj naučno-tekhnicheskoi konferencii «Otkrytie semanticheskie tekhnologii proektirovaniya intellektual'nyĭ sistem» (OSTIS-2012). Minsk, BGUIR, 16-18 fevralya 2012 g. S. 397–400.
- [13] Ÿzbek tilining imlo lufati. (Tuzuvchilar: E.A.Begmatov, A.P.Madvaliev; N.Maxmudov taxriri ostida). Toshkent: Akademnashr, 2013. 520 b.
- [14] Ÿzbek tilining izoxli lufati. I zhild. Toshkent: ŸzME, 2006. B.257.
- [15] Ÿzbek tilining izoxli lufati. II zhild. Toshkent: ŸzME, 2006. B.187-188.