

Джусупова У.Т.
Узбекский государственный университет мировых языков
Ташкент
E-mail: unglasin@mail.ru

ЛИНГВОСТАТИСТИКА: ПОНЯТИЯ, НАПРАВЛЕНИЯ, ПРИМЕНЕНИЕ

Аннотация. В статье рассматривается проблема лингвостатистики как сформировавшегося и интенсивно развивающегося направления в современном языкознании. Анализируются и описываются понятия и направления, а также методы и приемы математической лингвистики, которые используются как в целом в филологии, так и в языкознании, в частности. Отмечается, что междисциплинарность терминологического аппарата современных научных направлений способствует зарождению и формированию новых тенденций в науке о языке, что в целом обогащает филологическую науку как полинаучную систему. В этом контексте в статье осуществлен лингвостатистический анализ сингармонических и несингармонических односложных слов узбекского языка на материале двухтомного Толкового словаря узбекского языка. Выявлено, что основной массив односложных слов узбекского языка тюркского происхождения является сингармоническим. Часть односложных слов, заимствованных из фарси, арабского и русского языков или вошедших в узбекский язык через русский язык, в какой-то степени сингармонизирована. В современном узбекском литературном языке вследствие влияния фарси наблюдается нарушение сингармонизма в процессе аффиксации, когда, например, к сингармоническому корню слова присоединяется несингармонический аффикс.

Ключевые слова: лингвостатистика, математические методы, направления, частотность, теория вероятностей, односложное слово, сингармонизм, заимствования, тюркский, узбекский.

Dzhusupova U.T.
Uzbekistan State University of World Languages
Tashkent
E-mail: unglasin@mail.ru

LINGUISTIC STATISTICS: CONCEPTS, TRENDS, APPLICATIONS

Abstract. The article deals with the problem of linguistic statistics as a developed and intensively developing direction in modern linguistics. The concepts and directions, as well as the methods and techniques of mathematical linguistics, which are used both in general in philology and in linguistics, in particular, are analyzed and described. It is noted that the interdisciplinary terminological apparatus of modern scientific areas contributes to the emergence and formation of new trends in the science of language, which in general enriches the philological science as a polinauchnuyu system. In this context, the article carried out a linguistic-statistical analysis of synharmonic and non-harmonic monosyllabic words of the Uzbek language on the material of the two-volume Explanatory Dictionary of the Uzbek language. It is revealed that the main array of monosyllabic words of the Uzbek language of Turkic origin is synharmonic. Some of the monosyllabic words borrowed from Farsi, Arabic and Russian, or included in the Uzbek language through the Russian language, are to some degree

synharmonized. In modern Uzbek literary language, due to the influence of Farsi, a violation of synharmonicity is observed in the process of affixation, when, for example, a non-harmonic affix is added to the synharmonic root of a word.

Key words: linguistic statistics, mathematical methods, trends, frequency, probability theory, monosyllabic word, synharmonism, borrowings, Russian, Turkic, Uzbek.

Вступление

Современная эпоха развития науки в целом и филологической науки, в частности, является полинаучно-функциональной, так как междисциплинарность внедряется почти во все отрасли как в теоретическом, так и в прикладном характере, поэтому зарождаются и формируются новые направления, одни из которых уже сформировались, другие находятся на стадии становления. М. Джусупов отмечает, что будущее науки – в автономизации отдельной отрасли каждой большой науки [9, с. 7].

В языковедческой науке во второй половине XX в. – начале XXI в. сформированы новые направления, которые приобрели статус самостоятельных наук как внутри филологии, так и внутри лингвистики.

Одним из таких направлений является математическая лингвистика, в нашем случае, лингвистическая статистика (лингвостатистика), использование методов и приёмов которой в различных отраслях филологии способствовало раскрытию многих нюансов, которые методами традиционной лингвистики не поддавались полному анализу и получению соответствующих научно обоснованных результатов.

Лингвистическая статистика как научное направление

Определение качественных и количественных характеристик сингармонического и несингармонического слова в узбекском, казахском, русском и других языках имеет большое теоретическое и практическое значение. Статистический анализ будет способствовать раскрытию как количественных, так и качественных особенностей слова в сопоставляемых языках.

Определение сингармонических и несингармонических свойств слова осуществляется математическими методами. Всякое количественное

(математическое, статистическое) исследование, прежде всего, необходимо для раскрытия качественных свойств языковой единицы (в данном случае односложного слова). Следовательно, количественные показатели свидетельствуют о наличии качественных характеристик тех или иных языковых явлений. Это означает, что количественные показатели должны подвергаться качественной интерпретации.

Математические методы в лингвистике стали применяться сравнительно недавно. Мнения ученых по данному вопросу разнятся. Большинство лингвистов полагает, что «язык лежит вне меры и числа» [20, с. 15]. **Некоторые** языковеды считают, что количественная методика вполне достаточна для собственно лингвистических исследований, другие находят такое предположение ошибочным [1; 14; 2; 19].

В научных работах при анализе языковых единиц какого-либо типа, встречающихся в тексте, долгое время использовались слова «больше», «меньше», «мало», «много», «редко», «часто» и т.п. При таком подходе, когда игнорируются точные статистические данные, трудно установить строгие закономерности, действующие в языке и проявляющиеся в речевой деятельности.

Как было отмечено выше, язык и речь имеют не только качественные, но и количественные характеристики. Например, если обратиться к языковой системе, то можно подсчитать количество частей речи (самостоятельных и служебных), количество фонем в звуковой цепи (гласных и согласных) и т.д. В текстах (в речи) количественным показателем является частотность тех или иных языковых знаков, форм, словоупотреблений и т.д.

На необходимость применения математических методов в языкознании указывают не только лингвисты, но и математики. Так, еще в начале XX века А. А. Марков, проанализировав результаты работы Н.А. Морозова «Лингвистические спектры», пришел к выводу, что полученные результаты должны обрабатываться с применением теории вероятностей, при подсчетах

же нужно учитывать средние числа, размеры отклонений от средней величины [16, с. 30].

В настоящее время необходимость применения математических (в частности, статистических) методов в лингвистике не вызывает сомнений. В 90-е годы появились многочисленные работы по математической лингвистике, лингвостатистике [7; 22; 18; 17]. В известной книге Б.Н. Головина «Язык и статистика» [7] содержится систематическое изложение методики статистических исследований в языкознании. Р.М. Фрумкина считает, что конкретные исследования различных явлений языка и речи свидетельствуют об «очевидной органичности статистических методов для лингвистики как науки» [22].

Применение математических методов при изучении языковых явлений считается возможным, поскольку исследования анализируют два языка: язык математики и человеческий (естественный) язык. Р.Г. Пиотровский по этому поводу отмечает, что принципиальное сходство между этими двумя языками заключается в том, что как естественный язык, так и язык математики представляют собой знаковые (семиотические) системы, служащие для передачи информации [18, с. 5]. Вместе с этим следует учитывать и существенные расхождения между структурой языкового знака и знака математического. Например, расхождения, заключающиеся в том, что математическое множество имеет четкие границы: значение знака 9 представляет собой множество, которое включает в себя только девять (не восемь и не десять) элементов. Десигнативное же значение языкового знака можно рассматривать как множество денотатов, но это множество не всегда имеет четкие границы. Иначе говоря, значения языковых знаков представляют собой нечеткие множества с размытыми границами [18, с. 6].

Применяя математические методы, математическую статистику, следует нечеткие множества максимально приблизить к четким, сознательно отделив от второстепенных, несущественных признаков языковых знаков.

Математическая статистика возникла на основе теории вероятностей, с которой связана непосредственным и естественным образом. Теория вероятностей изучает события, которые могут происходить или не происходить, а если происходят, то с различной степенью вероятности. В практике опыта (испытаний) осуществляются события различных типов, в том числе и лингвистические события (следование фонемы или группы фонем после данной фонемы, следование слова или словосочетания после данного слова и др.) «Поскольку отдельные элементы языка в речи характеризуются некоторой частотой, а, следовательно, и вероятностью, произведения речи, очевидно, могут анализироваться и исследоваться с помощью методов математической статистики» [17, с. 5]. Очень часто различают «неколичественную» и «количественную» математику. К разделам «неколичественной математики» относят теорию множеств, математическую логику, теорию алгоритмов и др., которые находят свое применение в комбинаторной лингвистике, а к разделам «количественной математики» – математическую статистику, теорию вероятностей, теорию информации, математический анализ и др., которые легли в основу квантитативной (количественной) лингвистики [18, с. 8].

Если признать, что в языковой системе (структуре) имеются ограниченные (конечные) множества фигур – незначков (фонем, слогов, морфем, лексем, моделей словосочетаний и предложений), то следует согласиться с тем, что при изучении языковой системы (структуры) должны применяться методы комбинаторной лингвистики. В то же время система языка представляет собой механизм, порождающий тексты. В тексте (в речи) на встречаемость отдельных лингвистических единиц накладываются различные ограничения вероятностного характера. Иными словами, частотность лингвистических единиц различного типа и их сочетаний носит вероятностный характер. При исследовании явлений речи в наибольшей степени должны применяться методы квантитативной (количественной) лингвистики, с помощью которых исследуются «акустико-физиологическая и

психолингвистическая проблематика речеобразования, а также стилистика и история языка. При их изучении широко применяются не столько комбинаторные, сколько количественные измерения» [18].

Таким образом, квантитативная лингвистика имеет дело прежде всего со статистическими методами, на необходимость применения которых указывал еще В.В. Виноградов [6, с. 155].

К типично статистическим наукам относит лингвистику и языковед Пьер Гиро, так как «язык, по сути дела, является статистическим явлением, т.е. таким, которое охватывает измеряемые элементы, проявляющие статистические тенденции, а поэтому они могут быть определены и интерпретированы численно» [17, с. 21]. Следовательно, существует особая языковая функция, которую Гиро называет «статистической функцией», являющейся не менее важной и не менее объективной, чем функция семантическая, и столь же необходимой для полного и правильного понимания языковых фактов [17, с.21].

В лингвостатистике различают три основные направления:

«К первому и наиболее популярному среди лингвистов относят получение количественных данных о языке и речи в результате подсчета частот» [17, с. 23]. В рамках данного направления выполнен ряд работ [5; 8; 15; 11]. В статье А.А. Вейлерт «О некоторых факторах, определяющих частоту слова в тексте» [15, с. 99-103] была поставлена задача – выявить экстралингвистические и собственно лингвистические факторы, влияющие на частоту слова. Подсчету подлежали этимоны (этимологические единицы) в верхненемецком языке. В результате анализа в 32.092 (СУ) было обнаружено 1157 этимонов, подсчитана частотность словоупотреблений этимонов по историческим периодам. На основании этих подсчетов без применения специального аппарата статистики автор делает вывод о том, что «обнаруживается прямая и несомненная связь между «возрастом» этимологической единицы и ее частотой: чем эта единица старше, тем и частотнее» [5, с.100].

Более совершенную методику статистических исследований с применением матричных таблиц использует М.Ж. Зулхонов в своем диссертационном исследовании [15]. А.Х. Джубанов считает, что статистика имеет неоценимое значение в выявлении лингвистических особенностей языка, так как она помогает определить степень загруженности языковой системы теми или иными элементами [8, с. 12-21].

Как известно, количество фактически встречающихся сочетаний в языке меньше теоретически возможного количества, а отношения между ними представляют относительную частотность сочетаний любого структурного типа или конкретного сочетания вычисляется по формуле, в которой d – частотность сочетаний определенного типа, m – фактическая сочетаемость, n – теоретически возможная сочетаемость [23, с. 80].

М. Джусупов вводит понятие величина запрета в языке и речи и обозначает ее символом l . Это «количество звукосочетаний определенного типа невозможных в языке в силу существующих запретов на сочетаемость звуков (глухих со звонкими, звонких с глухими, свистящих перед шипящими...», то есть модернизирует формулу, которая в новом виде выглядит следующим образом $d = \frac{m}{n^2 - l}$ [11, с.78]. Согласно ей, «...из теоретически возможного количества фонемосочетаний вычисляется количество невозможных в данном структурном типе звукосочетаний и остается фактически возможное количество звукосочетаний.

М. Джусупов, применяя в своем исследовании данную формулу, отмечает, что ее помощи «можно вывести относительную частотность каждого типа фонем и звукосочетаний в языке – двух-, трех-, четырех- и пятиэлементных. Когда подсчитывается встречаемость структурных типов трехкомпонентных сочетаний, величина n берется в кубе (n^3), в четырехэлементных в квадратных (n^4)» [11, с.78]. По мнению автора, подобное решение распространяется только на консонансы однородных структурных типов, например, ПЯ+ПЯ (переднеязычный согласный +

переднеязычный согласный), ПЯ +ПЯ+ПЯ и др. Формула удобна для подсчета относительной частотности фонемосочетаний, т.е. на фонологическом уровне [11, с.79]. Качественные и количественные результаты анализа совпадают. Но различные типы текстов имеют разные частотные значения. Например, консонансы [лн] – вольный, больной и т.д. в общественно-политических текстах встретилось 85 раз, в драматургических – 12 раз, в научно-популярных текстах – 43 раза, а в текстах художественной литературы – 21 раз [11, с.79]. Такая поляризованность в результатах статистического исследования свидетельствует о том, что для изучения и обучения необходим анализ и словарей, и текстов.

Второму направлению лингвостатистики «принадлежит применение теории вероятностей для построения различных языковых моделей, при этом также используются данные, полученные в результате подсчета частот [17, с.23].

К третьему направлению относятся исследования, в которых «аппарат математической статистики применяется во всем его объеме для описания и обоснования поведения языковых элементов в языке и в речи по результатам исследований выборок с заранее установленной надежностью выводов [17, с. 23]. В его рамках широко используется структурно-вероятностный анализ (СВА) языка по данным речи, принципы которого разработаны Н.Д. Андреевым [3]. Метод СВА дает возможность объективно (независимо от позиции исследователя) описывать функционирование лингвистических элементов в любом языке, в том числе и в сопоставительном плане. К основным понятиям СВА относятся независимая вероятность (НВ) – вероятность лингвистических событий, не связанная с каким-либо условием; условная вероятность (УВ) – вероятность лингвистических событий, связанная с определенными условиями; коррелятивная функция (КФ) – частное, получаемое от деления величины УВ на величину НВ; абсолютное квадратическое отклонение (АКО); относительное квадратическое отклонение (ОКО); валентность (избирательная сочетаемость языковых

элементов) [4]. Пользуясь методом СВА, А.И. Бахарев пришел к заключению: статистические данные приводят к выводам теоретического характера, к результатам, которые нуждаются в собственно лингвистическом истолковании [4].

Лингвостатистический анализ сингармонических и несингармонических односложных слов узбекского языка

В работе применены методы комбинаторной лингвистики для лингвостатистического исследования односложных слов сингармонического и несингармонического характера на материале двухтомного Толкового словаря узбекского языка (1981). Предметом исследования являются словарные (языковые) единицы – односложные слова. Если слог (односложное слово) рассматривать как элементы множеств и подмножеств, то окажется, что эти подмножества и множества конечны. Например, «Толковой словарь узбекского языка» («Ўзбек тилининг изоҳли луғати». М.: Рус тили, 1981. 60 000 тыс. слов).

Лингвостатистический анализ производится с дифференциацией односложных слов в словаре на 5 групп (типов) по признаку сингармонизма в исконно тюркских узбекских и заимствованных словарных составах:

- 1) односложные сингармонические слова тюркского (узбекского) происхождения;
- 2) односложные сингармонизированные слова, заимствованные из фарси;
- 3) односложные сингармонизированные слова, заимствованные из арабского языка;
- 4) односложные слова, заимствованные из русского языка или через русский язык сингармонического характера;
- 5) односложные слова, заимствованные из русского языка или через русский язык несингармонического характера.

Результаты сравнительного статистического анализа внутри групп односложных слов по всему массиву словаря дали следующие результаты.

Всего односложных слов в словаре 1587, что составляет 100% всех односложных слов в словаре. Наиболее многочисленными и высокочастотными в узбекском языке являются односложные **сингармонические слова тюркского происхождения** (I тип – 702; то есть 44,2%), затем идут менее частотные заимствованные односложные слова сингармонизированные, заимствованные из фарси (II тип – 329; то есть 20,7%) и арабского языка (III тип – 300; то есть 18,9%), третьим по частотности являются заимствованные односложные слова из русского языка или через русский язык с сингармоническим маркером (IV тип – 171; то есть 10,7%), наиболее малочисленным являются заимствованные односложные слова из русского языка или через русский язык с несингармоническим произношением и написанием (ломанные, смешанные) (V тип – 85; то есть 5,4%).

Статистические данные, полученные в результате определения количественного соотношения между выделенными выше типами односложных слов узбекского языка, свидетельствуют о том, что исконно-узбекские односложные слова тюркского происхождения с сингармоническим произношением (I тип) имеют значительный удельный вес в словарном составе, точнее, почти $2/3$ всех односложных слов. Значительное место в словаре занимают заимствованные односложные слова с сингармоническим произношением (II, III, IV типы), на которые приходится примерно вторая часть всех односложных слов в словаре узбекского языка. И только восемнадцатая часть – на заимствованные односложные слова из русского языка или через русский язык с несингармоническим произношением и написанием (V тип).

Итак, проведенное статистическое исследование типов односложных сингармослов узбекского языка на материале толкового словаря показало, что в составе словаря преобладают односложные слова, характеризующиеся моносингармотембральностью (94, 5%). К ним мы относим заимствованные односложные слова из фарси, арабского языка русского языка или через русский язык с сингармоническим произношением (50, 3%). В группу полисингармонических (несингармонических) односложных слов мы

включили односложные слова, заимствованные из русского языка или через русский язык с несингармоническим произношением и написанием (5,4%). Они преимущественно используются в научной, публицистической, официально-деловой речи.

Таблица № 1.

Статистические характеристики односложных слов с сингармонической и несингармонической просодикой в узбекском языке на материале «Толкового словаря узбекского языка» («Ўзбек тилининг изоҳли луғати») [1981].

	Типы	В абсолютных числах	В процентах %
I.	Исконно-узбекские (тюркские) односложные слова сингармоническим произношением	702	44,2
II.	Заимствованные односложные слова из форси сингармоническим произношением	329	20,7
III.	Заимствованные односложные слова из арабского языка сингармоническим произношением	300	18,9
IV.	Заимствованные односложные слова из русского языка или через русский язык сингармоническим произношением	171	10,7
V.	Заимствованные односложные слова из русского языка или через русский язык несингармоническим произношением и написанием (ломанные, смешанные)	85	5,4%

Итак, 94,5% односложных слов узбекского языка в исследуемом словаре являются сингармоническими. При присоединении к ним, например, аффикса

множественного числа, все 94% сингармонических односложных слов становятся несингармоническими (под влиянием фарси). Например, бош+лар [бошләр], тоғ [тоғләр], кыр [кырләр] и т.д., что стало нормой узбекского литературного языка, хотя в нем серьезно нарушен принцип сингармонизма – основы всей тюркской звучащей речи [12; 10; 13].

Заключение

Применение математических методов и приемов в филологостатистических и в лингвостатистических целях имеет большое теоретическое и практическое значение. Статистика в целом в жизни общества и науковедения и лингвостатистика как ее ответвление, специальное направление в филологической науке, работает на определение и анализ, прежде всего, качественных свойств языковых единиц и их сочетаний как в парадигматическом, так и в синтагматическом аспектах, ибо количественные (статистические) показатели характеризуют не столько «арифметические» свойства, сколько их (в большой степени) смысловые, семантико-стилистические, словообразовательные, фонетико-фонологические, морфологические, синтаксические возможности как в аспекте потенциального, так и в аспекте фактического функционирования в различных типах текстов и словарей.

Использование статистических и приемов, используемых при исследовании языка и речи, способствуют углубленному и широкому раскрытию внутренней сути языка как средства общения, сообщения и воздействия. Результаты таких исследований актуальны и в плане осуществления фундаментальных научных изысканий, и в плане создания учебных материалов нового поколения в эпоху параллельного функционирования сравнительно-исторической, системно-структурной и антропоцентрической парадигмы научных мнений.

БИБЛИОГРАФИЯ

1. Адмони В.Г. Качественный и количественный анализ грамматических явлений // Вопросы языкознания. 1986. № 6. С.17 – 20.
2. Алефиренко Н.Ф. Современные проблемы науки о языке: Учебное пособие. М: Флинта: Наука, 2005. 416 с.
3. Андреев Н.Д. Статистико-комбинаторные методы в теоретическом и прикладном языкознании. Л.: Наука, 1967. 403 с.

4. Бахарев А.И. Отрицание в логике и грамматике. Саратов: Изд-во Саратовского государственного университета, 1980. 77 с.
5. Вейлерт А.А. О некоторых факторах, определяющих частоту слова в тексте // Вопросы языкознания. 1978. № 2. С.99-103.
6. Виноградов В.В. Современный русский язык. М.: Учпедгиз, 1938. Вып.1. 160 с.
7. Головин Б.Н. Язык и статистика. М.: Просвещение, 1971. 190 с.
8. Джубанов А.Х. У истоков идей математизации лингвистических исследований // Статистика казахского текста. Алма-Ата: Гылым, 1990. С.12 – 21.
9. Джусупов М. Автономизация науки (лингводидактика и методика): Материалы международной научно-практической конференции «Теория и практика преподавания языка и литературы: современное состояние, проблемы, перспективы». Алматы, 2008. С.7 – 9.
10. Джусупов М. Ахмет Байтұрсынов және қазіргі қазақ тілі фонологиясы. Алматы: Гылым, 1998. 214 б.
11. Джусупов М. Звуковые системы русского и казахского языков. Слог. Интерференция. Обучение произношению. Ташкент, «Фан», 1991, 240 с.
12. Джусупов М. Фонемография А. Байтұрсынова и фонология сингармонизма. Ташкент: Узбекистан, 1995.
13. Дустанова У.Т. Сингармонизм как просодическое свойство слова (на материале односложного слова казахского, узбекского и русского языков) // Вестник КазНУ. Серия филологическая. №1 (91), 2006. С. 55 – 59.
14. Зиндер Л.Р., Строева Т.В. К вопросу о применении статистики в языковедении // Вопросы языкознания. 1968. № 6. С.25 – 29.
15. Зулхонов М.Ж. Лингвостатистический анализ конечных двухэлементных консонантных сочетаний в разносистемных языках (на материале немецкого и узбекского языков): Автореф. дис. канд. филол. наук. Ташкент, 1998. 21 с.
16. Марков А.А. Об одном применении статистических методов // Известия АН СССР. Сер.6. 1916. № 4. С. 30 – 33.
17. Носенко И.А. Начало статистики для лингвистов. М.: Высшая школа, 1981. 157 с.
18. Пиотровский Р.Г., Бектаев К.Б., Пиотровская А.А. Математическая лингвистика. М.: Высшая школа, 1977. 383 с.
19. Сидоренко. Е.В. Методы математической обработки в психологии. Санкт-Петербург: Речь, 2010. 350 с.
20. Трубецкой Н.С. Основы фонологии. М.: Изд-во иностранной литературы, 1960. 372 с.
21. Ўзбек тилининг изохли луғати. Икки томли. Маъруфов З.М. (ред.) М.: Русский язык, 1981. 632 с.
22. Фрумкина Р.М. Роль статистических методов в современных лингвистических исследованиях // Математическая лингвистика: сб. научн. тр. М.: Высшая школа, 1973. С. 57 – 60.
23. Шаброва В.А. Опыт сопоставительного изучения консонантных сочетаний // Русский язык в национальной школе. 1977. № 4. С. 80 – 82.

REFERENCES

1. Admoni V.G. Kachestvennyj i kolichestvennyj analiz grammaticheskikh yavlenij // Voprosy yazykoznanija. 1986. №6. S. 17 – 20.
2. Alefirenko N.F. Sovremennye problemy nauki o yazyke: Uchebnoe posobie. M: Flinta: Nauka, 2005. 416 s.
3. Andreev N.D. Statistiko-kombinatornye metody v teoreticheskom i prikladnom yazykoznanii. L.: Nauka, 1967. 403 s.

4. Baharev A.I. Otricanie v logike i grammatike. Saratov: Izd-vo Saratovskogo gosuniversiteta, 1980. 77 s.
5. Vejlert A.A. O nekotoryh faktorah, opredelyayushchih chastotu slova v tekste // Voprosy yazykoznanija. 1978. № 2. S. 99 –103.
6. Vinogradov V.V. Sovremennyy russkij yazyk. M.: Uchpedgiz, 1938, vyp.1. 160 s.
7. Golovin B.N. YAzyk i statistika. M.: Prosveshchenie, 1971. 190 s.
8. Dzhubanov A.H. U istokov idej matematizacii lingvisticheskikh issledovanij // Statistika kazahskogo teksta. Alma-Ata: Gylym, 1990. S.12–21.
9. Dzhusupov M. Avtonomizaciya nauki (lingvodidaktika i metodika): Materialy mezhdunarodnoj nauchno-prakticheskoj konferencii «Teoriya i praktika prepodavaniya yazyka i literatury: sovremennoe sostoyanie, problemy, perspektivy». Almaty, 2008. S.7–9.
10. Dzhusupov M. Ahmet Bajtyrsynov zhəne qazirgi qazaq tili fonologiyasy. Almaty: Fylym, 1998. 214 b.
11. Dzhusupov M. Zvukovye sistemy russkogo i kazahskogo yazykov. Slog. Interferenciya. Obuchenie proiznosheniya. Tashkent, «Fan»,1991, 240 s.
12. Dzhusupov M. Fonemografiya A. Bajtursynova i fonologiya singlarmonizma. Tashkent: Uzbekistan, 1995.
13. Dustanova U.T. Singarmonizm kak prosodicheskoe svojstvo slova (na materiale odnoslozhnogo slova kazahskogo, uzbekskogo i russkogo yazykov)// Vestnik KazNU. Seriya filologicheskaya. №1 (91), 2006. S. 55 – 59.
14. Zinder L.R., Stroeva T.V. K voprosu o primenenii statistiki v yazykovedenii // Voprosy yazykoznanija. 1968. № 6. S. 25 – 29.
15. Zulhonov M.Zh. Lingvostatisticheskij analiz konechnykh dvuhelementnykh konsonantnykh sochetanij v raznosistemnykh yazykah (na materiale nemeckogo i uzbekskogo yazykov): Avtoref. dis. kand. filol. nauk. Tashkent, 1998. 21 s.
16. Markov A.A. Ob odnom primenenii statisticheskikh metodov // Izvestiya AN SSSR. Ser.6. 1916. № 4. S. 30 – 33.
17. Nosenko I.A. Nachalo statistiki dlya lingvistov. M.: Vysshaya shkola, 1981. 157 s.
18. Piotrovskij R.G., Bektaev K.B., Piotrovskaya A.A. Matematicheskaya lingvistika. M.: Vysshaya shkola, 1977. 383 s.
19. Sidorenko. E.V. Metody matematicheskoy obrabotki v psihologii. Sankt-Peterburg: Rech', 2010. 350 s.
20. Trubeckoj N.S. Osnovy fonologii. M.: Izd-vo inostrannoj litereratury, 1960. 372 s.
21. Ÿzbek tilining izoxli lu'fati. Ikki tomli. Ma"rufov Z.M. (red.) M.: Russkij yazyk, 1981. 632 s.
22. Frumkina R.M. Rol' statisticheskikh metodov v sovremennykh lingvisticheskikh issledovaniyah // Matematicheskaya lingvistika: sb. nauchn. tr. M.: Vysshaya shkola, 1973. S.57 – 60.
23. Shabrova V.A. Opyt sopostavitel'nogo izuchenii konsonantnykh sochetanij // Russkij yazyk v nacional'noj shkole. 1977. № 4. S. 80 – 82.