

УДК 8.811.512.14

Статистическая оценка фонологических данных тюркских языков в современном языкознании

Аврутина Аполлиария Сергеевна

Кандидат филологических наук, доцент,
кафедра теории и методики преподавания языков и культур Азии и Африки,
Санкт-Петербургский государственный университет,
190000, Российская Федерация, Санкт-Петербург, Университетская наб., 11;
e-mail: a.avrutina@spbu.ru

Аннотация

Настоящая статья посвящена обзору опытов количественного анализа фонологических и морфонологических элементов в текстах на агглютинативных (тюркских) языках в специальной литературе. В языкознании первые работы, посвященные непосредственно количественному исследованию фонологической подсистемы различных языков, появились в середине XIX века начиная со случаев применения математических методов исследования в работах представителей казанской лингвистической школы. Работ, посвященных частотности употребления фонологических единиц на материалах различных языков, существует немало, однако на материалах тюркских языков, и особенно на фонологическом материале, метод количественного анализа до сих пор применялся сравнительно редко. В настоящее время в специальной литературе не выработано единой методики статистического исследования материала агглютинативных (тюркских) языков. Выбор объектов и критериев исследования каждый раз проводится автором независимо, вследствие чего сравнение полученных в различных работах результатов затруднено. Сравнение различных работ показывает, что анализ какой-либо одной подсистемы языка не показателен. Кроме того, в различных исследованиях используется различный объем текстовой выборки, который варьируется от небольшого фрагмента на несколько сотен слов до целого корпуса на несколько миллионов слов. Различные объемы выборки в работах различных исследователей дают неодинаковые показатели на материале даже одного языка. Статистическое исследование зависит от многих факторов, в частности от стиля выбранного текста. Работа по вычислению типологических индексов должна строиться на вычислении множества значений различных типов индексов и на установлении разницы между ними.

Для цитирования в научных исследованиях

Аврutiна А.С. Статистическая оценка фонологических данных тюркских языков в современном языкознании // Культура и цивилизация. 2016. Том 6. № 5А. С. 29-39.

Ключевые слова

Агглютинативные языки, тюркские языки, квантитативная лингвистика, фонология, тюркская фонология, тюркская морфонология.

Введение

Методы количественного анализа текстов отчасти применялись грамматологами в процессе исследования древних письменных памятников с целью реконструкции мертвых языков и в процессе дешифровки текстов [Сегал, 1972, 19]. Например, Жан-Франсуа Шампольон, который продолжил после Томаса Юнга и завершил дешифровку египетской системы письма в 1822 году чтением Розеттской надписи, обратил внимание на частоту повторяемости отдельных рисуночных знаков (а следовательно, слов) [Гельб, 2004]. Таким образом, работу по дешифровке древнетюркских рунических надписей, выполненную 25 ноября 1893 году датский ученым В. Томсеном, также можно условно считать первой попыткой количественного исследования текста на тюркском языке [Кононов, 1980, 3-9; Сегал, 1972, 20-21].

Обзор количественных исследований фонологической подсистемы различных языков

В языкознании первые работы, посвященные непосредственно количественному исследованию фонологической подсистемы различных языков, появились в середине XIX века и позже. Случаи применения математических методов исследования наблюдаются в работах представителей казанской лингвистической школы. Например, в «Общем курсе грамматики русского языка» В.А. Богородицкого предпринимается попытка квантитативной оценки фонемного состава русского языка [Богородицкий, 2005, 39-40].

Принципиально нового уровня квантитативный анализ фонологической системы достигает в 30-е годы XX века в рамках исследований Пражского лингвистического кружка. Например, Б. Трнка занимался статистическими подсчетами на материале фонологической системы английского языка [Trnka, 1966], В.Ф. Тводделл подсчитывал согласные фонемы и их сочетания в немецком языке [Twaddell, 1938], Д.К. Ципф окончательно сформулировал основные положения фонологической статистики [Zipf, 1932], он, по сути, является основателем современной квантитативной лингвистики.

Метод «фонологической статистики» был упомянут Н.С. Трубецким в «Основах фонологии», хотя именно Н.С. Трубецкому принадлежит высказывание о том, что язык, прежде

всего, – это система отношений, а поэтому «...лежит вне меры и числа» [Трубецкой, 2000, 16]. При этом в «Основах фонологии» Н.С. Трубецкой отметил, что статистические методы могут находить применение в фонологических исследованиях, когда предметом исследования является сочетаемость фонем, при этом он отмечает, что «подсчитывать следует не буквы и звуки, а фонемы и сочетания фонем» [там же, 272]. Подобная методика отчасти превзошла идеи будущих структуралистов. Н.С. Трубецкой называет два возможных метода подсчета фонологических показателей. С одной стороны, целью этого подсчета должно стать выявление частотности определенных фонологических элементов в рассматриваемом языке (фонем, сочетаний фонем, определенных типов слов и морфем), с другой стороны, желательно подсчитывать степень загруженности рассматриваемых фонологических элементов и их фонологических оппозиций [там же, 272-273].

Опыт исследования частоты встречаемости индивидуальных фонем был предпринят в работе Е. Черри, М. Халле, Р. Якобсона [Черри, Халле, Якобсон, 1962]. Авторы подчеркивают, что воспользовались методами статистической теории информации. Слова в речи понимаются авторами данной статьи как единицы кода. Определение минимального набора дифференциальных признаков предполагает выделение избыточных фонетических различий. Язык понимается ими как математический процесс. Для выявления частотности фонем авторами было выбрано три типа подсчетов: 1) подсчет, который учитывает границы слов и стыки между соседними частями сложных слов, обозначаемые дефисом; 2) подсчет, который учитывает только границы слов; 3) подсчет, который не учитывает ни границы слов, ни стыки сложных слов, а учитывает только перерыв последовательности фонем в местах обязательной паузы [там же, 282-283]. Авторы отмечают, что статистический анализ фонем и их последовательностей в связных сообщениях необходимо дополнить аналогичным анализом словаря, который необходим для понимания дистрибуции фонем в лексическом коде рассматриваемого языка [там же, 297].

Продолжая идеи Н.С. Трубецкого, Р. Якобсон, Г.М. Фант и М. Халле создали так называемую дихотомическую (двоичную) описательную систему, содержащую минимальное число бинарных единиц, которая позволяла описывать в двоичных параметрах подсистемы фонем различных языков [Якобсон, Фант, Халле, 1962].

Среди последователей системы пражской школы, а также Р. Якобсона и М. Хале в отечественной лингвистике следует особо упомянуть лингвистов Л.Г. Зубкову и В.И. Золхоева, работавших с тюркским материалом.

В работе Л.Г. Зубковой «Фонологическая типология слова» была предпринята попытка статистического исследования вокалической и консонантной структуры тюркского слова [Зубкова, 1990]. Анализируя положение гласных и согласных фонем в простых и многоморфемных словах различных языков, она высчитывает количественную дистрибуцию фонем в слове согласно их качественным характеристикам и приводит данные по нескольким языкам, в частности турецкому. Например, степень дифференциации фонем по их дистрибуционной активности в турецком языке составляет 0,773 в начале, 0,500 в середине и 0,727 в конце слова [там же, 43].

В.И. Золхоев составил справочник-классификацию фонем различных языков мира посредством цифрового кода. Среди прочих В.И. Золхоев предпринял попытку рассмотреть дифференциальные признаки гласных и согласных нескольких тюркских языков методом количественной и статистической оценки. Развивая указанный подход, он выявил состав и количество дифференциальных признаков фонем, а также состав и количество бинарных оппозиций в нескольких тюркских языках. При этом под системой оппозиций В.И. Золхоев понимает противопоставление каждой фонемы всем остальным фонемам внутри системы по составу их дифференциальных признаков. В монографии «Классификация фонем посредством цифрового кода» [Золхоев, 2006] В.И. Золхоев синтезирует методы пражской лингвистической школы, в частности Н.С. Трубецкого, развивая его теорию существенных признаков фонем и теорию оппозиций фонем, а также предлагает сравнить состав и количество бинарных оппозиций в различных языках мира, применяя специальный цифровой код [там же, т. 1, 5].

Большим вкладом в квантитативные исследования языков мира (правда, в сфере морфологии) стали работы Дж. Гринберга [Greenberg, 1960], целью которого было выявить те или иные типологические характеристики языков и который доработал и формализовал типологические постулаты Э. Сепира [Сепир, 1993]. Э. Сепир выдвинул положение о том, что языков «чистого» строя не существует – чисто агглютинативных, чисто флективных и т. д.; в каждом языке можно обнаружить элементы флективности, изоляции, агглютинативности и т. п. Развивая эти идеи, Дж. Гринберг писал: «...Язык вполне может содержать и обычно действительно содержит как агглютинирующие, так и неагглютинирующие конструкции» [Greenberg, 1960, 10]. В 1960 году Дж. Гринберг первым предложил оценивать тип языка с помощью квантитативно-типологических индексов [там же, 15]. Каждый из этих индексов должен был представлять собой дробь, числителем которой выступало число, например, префиксов в некотором тексте, а знаменателем – число слов, т. е. PREF/W есть индекс префиксальности для данного языка, где PREF – число префиксов, а W – число слов. Соответствующие величины подсчитывались на материале 100-словных текстов для ряда языков [Касевич, Аврутина, 2014].

К сожалению, в методике Дж. Гринберга отсутствовали как четкие указания по стилистике выбранного текста, так и строго универсальный понятийно-терминологический аппарат, в рамках которого следует подсчитывать те или иные языковые единицы. В то же время типологические индексы, выработанные Дж. Гринбергом, больше касаются, как уже было сказано выше, морфологии и синтаксиса, нежели фонологии.

Модификация и дополнение существующих индексов Дж. Гринберга была выполнена также в коллективной монографии «Квантитативная типология языков Азии и Африки» 1982 года под редакцией В.Б. Касевича и С.Е. Яхонтова [Касевич, Яхонтов, 1982]. В частности, был учтен важнейший недостаток типологии Гринберга – неразличение знаменательных и служебных слов [там же, 37]. При этом были введены 14 индексов, которые просчитываются для 26 восточных языков различного типа и ареальной принадлежности.

Один из разделов этой монографии, написанный ленинградским тюркологом В.Г. Кондратьевым, был посвящен турецкому языку. Для исследования был взят фрагмент текста в 100 слов, в котором разграничены морфемы, а также приводится подстрочный перевод. Затем приводится литературный перевод фрагмента, а далее следует таблица с типологическими индексами и комментарии [там же, 152-159]. В завершении составители монографии приходят к выводу, что агглютинативные языки обладают средним показателем агглютинации, высоким показателем словоизменения, управления, низким показателем примыкания и аналитичности, а также средним показателем лексичности [там же, 320-322].

Результаты, полученные в 1982 году, были радикально пересмотрены и дополнены в следующей коллективной монографии по данному вопросу, изданной уже в 2011 году, также под редакцией В.Б. Касевича, «Грамматика и семантика восточного текста. Квантитативные характеристики» [Касевич, 2011]. В этой монографии предлагается уже 12 типологических индексов для 14 восточных языков разных типов.

В указанной монографии не были представлены тюркские языки, которые были проанализированы по сходному набору индексов в рамках данного подхода позднее, в статье В.Б. Касевича и А.С. Аврутиной под названием «Грамматическая и статистическая структура текста на агглютинативном языке в исторической перспективе (на материале османского и турецких текстов)» [Касевич, Аврутина, 2014].

Следует упомянуть цикл работ Ю.А. Тамбовцева, посвященных квантитативному описанию различных систем языка. Особенностью работ Ю.А. Тамбовцева является стремление исследователя не пользоваться разработанными предшественниками квантитативно-типологическими индексами, а предложить некоторые собственные методы для подсчета фонологических показателей [Тамбовцев, 2010]. В своих работах он анализирует, в основном, типологию фонологических цепей в различных языках мира, уделяя немало внимания и тюркским языкам. Под «звуковой цепочкой» Ю.А. Тамбовцев понимает последовательность звуков, т. е. слова, которые в совокупности составляют текст [Тамбовцев, 2011, 109], (хотя уже этот термин вызывает большие сомнения и очевидно отвлекает читателя от собственно фонологических, языковых явлений, неверно направляя рассуждения в сферу речи – в сферу сиюминутных материальных образований). В статье «Фоно-метрические расстояния между азербайджанским и другими тюркскими языками с точки зрения типологии его звуковых цепочек» он сравнивает фоно-метрические расстояния между азербайджанским языком и другими тюркскими языками, сопоставляя фонологические цепочки различных тюркских языков и математически доказывая особое положение азербайджанского [Гаджиева, 1966].

В 2010 году в «Journal of Quantitative Linguistics» («Вестнике квантитативной лингвистики») была опубликована серия работ, в которых статистическому анализу был подвергнут лексический состав определенного среза того или иного языка. Одна из этих работ была посвящена анализу словарных характеристик турецкой литературы XX века [Fazli, Patton,

2010]. Авторы статьи, посвященной анализу лексики турецких романов, вышедших на протяжении XX века, задаются вопросом, как менялся лексический состав турецких романов на протяжении века. Для анализа были взяты 40 романов 40 различных авторов. Прделанный сравнительный анализ был основан на перекрестной проверке наиболее часто встречающихся слов (служебные слова – союзы, частицы, указательные местоимения и послелогии), которые в статье именуется дискриминаторами [там же, 167]. Авторы демонстрируют, что служебные слова употребляются чаще знаменательных, и статистико-стилистический анализ дает возможность с высокой точностью определять пол автора текста в зависимости от часто употребляемых им слов [там же, 180]. Кроме того, авторами установлено, что к концу XX века тексты стали сложнее, а слова в них стали использоваться более длинные [там же, 184]. Последний вывод, как представляется, имеет большую важность для работ по истории тюркских языков и позволяет сделать вывод о развитии турецкого языка в будущем. Авторы также вывели некоторые характеристики часто употребляемых слов в турецком языке XX века и доказали, что изменение общей длины слов в литературе конца XX века (общее удлинение слов) влияет на авторский выбор слов в каждом конкретном тексте [там же, 169].

Некоторое количество исследований, посвященных квантитативному анализу различных подсистем турецкого языка, было проведено в рамках конференции CICling. Хотелось бы остановиться на одном из них. Работа «Corpus frequency and affix ordering in Turkish» [Aksan, Demirhan, Aksan, www] посвящена квантитативному анализу корпуса турецкого языка, при котором исследуются мультиморфемные структуры в современном турецком языке. Авторы проводят статистический анализ комбинаций аффиксов на основании данных о частотности их употребления. Для исследования берутся стилистически различные тексты из TNC (Turkish National Corpus). Размер TNC – 50 997 016 употребляемых в настоящее время слов различных стилистических категорий, при этом тексты представляют период с 1990 по 2013 годы.

Интересное статистическое исследование было проведено в области турецкой фонетики. Автором этого исследования является турецкий лингвист Сюер Экер. Одна из его статей называется «Фонема /e/ как социофонетический индикатор». В работе проведен статистический анализ употребления аллофонов фонемы /e/ в различных областях Турции [Eker, 2007].

Заключение

Проведенный анализ работ, созданных в области статистических исследований на тюркском материале, убеждает в следующем:

1) в настоящее время в специальной литературе не выработано единой методики статистического исследования материала агглютинативных языков. Выбор объектов и критериев исследования каждый раз проводится автором независимо, вследствие чего сравнение полученных в различных работах результатов затруднено;

2) очевидно, что анализ какой-либо одной подсистемы языка не будет показателен;

3) в различных исследованиях используется различный объем выборки, который варьируется от небольшого фрагмента на несколько сотен слов до целого корпуса на несколько миллионов слов. Различные объемы выборки в работах различных исследователей дают неодинаковые показатели на материале одного языка;

4) статистическое исследование зависит от многих факторов, в частности от стиля выбранного текста.

Из всего вышесказанного следует, что работа по вычислению типологических индексов должна строиться на вычислении множества значений различных типов индексов и на установлении разницы между ними.

Библиография

1. Богородицкий В.А. Общий курс русской грамматики (из университетских чтений). М.: КомКнига, 2005. 576 с.
2. Гаджиева Н.З. Азербайджанский язык // Языки народов СССР. Т. 2. Тюркские Языки. М., 1966. С. 66-90.
3. Гельб И.Е. Опыт изучения письма. Основы грамматики. М.: Едиториал УРСС, 2004. 368 с.
4. Гринберг Дж. Квантитативный подход к морфологической типологии языков // Новое в лингвистике. 1963. Вып. 3. С. 60-94.
5. Золхоев В.И. Классификация фонем посредством цифрового кода (в 12 т.). Улан-Удэ: Изд-во Бурятского государственного университета, 2006.
6. Зубкова Л.Г. Фонологическая типология слова. М.: Университет дружбы народов, 1990. 254 с.
7. Касевич В.Б., Аврутина А.С. Грамматическая и статистическая структура текста на агглютинативном языке в исторической перспективе (на материале староосманских и турецких текстов) // Вестник Пермского университета. Российская и зарубежная филология. 2014. Т. 4. № 28. С. 7-14.
8. Касевич В.Б. (ред.). Грамматика и семантика восточного текста. Квантитативные характеристики. СПб.: СПбГУ, Восточный ф-т, Издательство РХГА, 2011. 213 с.
9. Касевич В.Б., Яхонтов С.Е. (ред.). Квантитативная типология языков Азии и Африки. Л., 1982. 350 с.
10. Кононов А.Н. Грамматика языка тюркских рунических памятников (VII – IX вв.). Л.: Наука, 1980. 256 с.
11. Сегал Д.М. Основы фонологической статистики (на материале польского языка). М.: Наука, 1972. 256 с.
12. Сепир Э. Избранные труды по языкознанию и культурологии. М.: Прогресс, 1993. 656 с.
13. Тамбовцев Ю.А. Фоно-метрические расстояния между азербайджанским и другими тюркскими языками с точки зрения типологии его звуковых цепочек // Rocznik Orientalistyczny. T. LXIV. Z. 2. 2011. С. 105-119.

14. Тамбовцев Ю.А. Фоно-типологическая близость лингвистических объектов по критерию «хи-квадрат» // Вестник НГУ. Серия: Информационные технологии. 2010. Т. 8. Вып. 3. С. 45-54.
15. Трубецкой Н.С. Основы фонологии. М.: Аспект-Пресс, 2000. 352 с.
16. Черри Е., Халле М., Якобсон Р. К вопросу о логическом описании языков в их фонологическом аспекте // Новое в лингвистике. Вып. 2. М.: Изд-во иностранной литературы, 1962. С. 279-298.
17. Якобсон Р.О., Фант Г.М., Халле М. Введение в анализ речи. Различительные признаки и их корреляты. Гл. II. Опыт описания различительных признаков // Новое в лингвистике. Вып. 2. М.: Изд-во иностранной литературы, 1962. С. 173-230.
18. Aksan M., Demirhan U.U., Aksan Y. Corpus frequency and affix ordering in Turkish. URL: <http://site.cicling.org/2016/accepted-abstracts.html>?
19. Eker S. Toplumdilbilgisel olarak /e/ // Edebiyat ve Dil Yazıları. Mustafa İsen'e Armağan. Ankara, 2007. P. 231-247.
20. Fazli C., Patton Jon M. Change of Word Characteristics in 20-Century Turkish Literature: A Statistical Analysis // Journal of Quantitative Linguistics. 2010. Vol. 17 (3). P. 167-190.
21. Greenberg J.H. A Quantitative Approach to the Morphological Typology of Language // International Journal of American Linguistics. 1960. Vol. 26. P. 178-194.
22. Trnka B. A Phonological Analysis of Present Day Standart English. Tokyo: Hokuon Publishing Go., 1966. 155 p.
23. Twaddell W.P. A phonological analysis of intervocalic consonant clusters in modern German // Actes du IV Congrès international de linguistes à Copenhague. 1938. P. 218-225.
24. Zipf G.K. Selected studies of the principle of relative frequency in language. Cambridge, Mass.: Harvard University Press. 1932.

Statistical evaluation of phonological data of Turkic languages in modern linguistics

Apollinariya S. Avrutina

PhD in Philology, Associate Professor,

Department of theory and methodology of teaching languages and cultures of Asia and Africa,

Saint Petersburg State University,

190000, 11 Universitetskaya emb., Saint Petersburg, Russian Federation;

e-mail: a.avrutina@spbu.ru

Abstract

This article reviews the experiences of quantitative analysis of phonological and morphological elements in the agglutinative (Turkic) texts. First works devoted to quantitative study of phonological subsystems of different languages appeared in the middle of XIX century, starting with the applications of mathematical methods in works of the representatives of the Kazan linguistic school. There is many works devoted to the usage frequency of phonological units on materials of different languages, but Turkic language materials, and especially quantitative analysis method of the phonological materials still are used relatively seldom. Today, the literature has not a single methodology of statistical research material for agglutinative (Turkic) languages. The selection of the objects and criteria in different researches is irrespective, so that the comparison of the results in various studies is difficult. Comparison of different works shows that the analysis of a single language subsystem is not indicative. In addition, various studies use text of various volume that varies from a small fragment for a few hundred words to the corpus for millions of words. Various works of various researchers give unequal performance on even the one language material. Statistical research depends on many factors, in particular, from the style of the selected text. The work on the calculation of the typological indices should be based on the calculation of different types of indexes and on the establishment of a difference between them.

For citation

Avrutina A.S. (2016) Statisticheskaya otsenka fonologicheskikh dannykh tyurkskikh yazykov v sovremennom yazykoznanii [Statistical evaluation of phonological data of Turkic languages in modern linguistics]. *Kul'tura i tsivilizatsiya* [Culture and Civilization], 6 (5A), pp. 29-39.

Keywords

Agglutinative languages, Turkic languages, quantitative linguistics, phonology, Turkic phonology, Turkic morphophonology.

References

1. Aksan M., Demirhan U.U., Aksan Y. Corpus frequency and affix ordering in Turkish. Available at: <http://site.cicling.org/2016/accepted-abstracts.html?> [Accessed 12/10/16].
2. Bogoroditskii V.A. (2005) *Obshchii kurs russkoi grammatiki (iz universitetskikh chtenii)* [General course of Russian grammar (from university readings)]. Moscow: KomKniga Publ.
3. Cherri E., Khalle M., Yakobson R. (1962) K voprosu o logicheskom opisaniy yazykov v ikh fonologicheskome aspekte [Toward the logical description of languages in their phonemic aspect]. In: *Novoe v lingvistike* [New horizons in linguistics], Vol. 2. Moscow, pp. 279-298.

4. Gadzhieva N.Z. (1966) Azerbaidzhanskii yazyk. T. 2. Tyurkskie Yazyki [Azerbaijani language. Vol 2: Turkic Languages]. In: *Yazyki narodov SSSR* [The languages of the peoples of the USSR]. Moscow, pp. 66-90.
5. Gel'b I.E. (2004) *Opyt izucheniya pis'ma. Osnovy grammatologii* [The experience of learning letters. Fundamentals of grammatology]. Moscow: Editorial URSS Publ.
6. Grinberg Dzh. (1963) Kvantitativnyi podkhod k morfologicheskoi tipologii yazykov [Quantitative approach to the morphological typology of languages]. *Novoe v lingvistike* [New horizons in linguistics], 3, pp. 60-94.
7. Kasevich V.B., Avrutina A.S. (2014) Grammaticheskaya i statisticheskaya struktura teksta na agglutinativnom yazyke v istoricheskoi perspektive (na materiale staroosmanskikh i turetskikh tekstov) [Grammatical and statistical structure of a text in the agglutinative language in historical perspective (basing on Turkish and Ottoman texts)]. *Vestnik Permskogo universiteta. Seriya "Rossiiskaya i zarubezhnaya filologiya"* [Perm University Herald. Series "Russian and Foreign Philology"], 28 (4), pp. 7-14.
8. Kasevich V.B. (ed.). (2011) *Grammatika i semantika vostochnogo teksta. Kvantitativnye kharakteristiki* [Grammar and semantics of the Eastern text. Quantitative characteristics]. Saint Petersburg: Saint Petersburg State University, Publishing house of Russian Christian Humanities Academy.
9. Kasevich V.B., Yakhontov S.E. (eds). (1982) *Kvantitativnaya tipologiya yazykov Azii i Afriki* [Quantitative typology of Asian and African languages]. Leningrad.
10. Kononov A.N. (1980) *Grammatika yazyka tyurkskikh runicheskikh pamyatnikov (VII – IX vv.)* [The grammar of the language of the Turkic runic inscriptions (7-9th centuries)]. Leningrad.
11. Segal D.M. (1972) *Osnovy fonologicheskoi statistiki (na materiale pol'skogo yazyka)* [The foundations of phonological statistics (on the material of Polish language)]. Moscow: Nauka Publ.
12. Sepir E. (1993) *Izbrannye trudy po yazykoznaniiyu i kul'turologii* [Selected works on linguistics and cultural studies]. Moscow: Progress Publ.
13. Tambovtsev Yu.A. (2010) Fono-tipologicheskaya blizost' lingvisticheskikh ob'ektov po kriteriyu "khi-kvadrat" [Phono-typological closeness of linguistic objects by the "chi-square" criterion]. *Vestnik NGU. Seriya "Informatsionnye tekhnologii"* [Herald of Novosibirsk State University. Series "Information technology"], 3 (8), pp. 45-54.
14. Tambovtsev Yu.A. (2011) Fono-metricheskie rasstoyaniya mezhdru azerbaidzhanskim i drugimi tyurkskimi yazykami s tochki zreniya tipologii ego zvukovykh tsepochek [Phono-metrical distances between the Old Turkic language and the other Turkic languages (based on the typology of its sound chains)]. *Rocznik Orientalistyczny*, 2 (64), pp. 105-119.
15. Trubetskoi N.S. (2000) *Osnovy fonologii* [Fundamentals of phonology]. Moscow: Aspekt-Press Publ.
16. Jakobson R.O., Fant G.M., Halle M. (1962) Vvedenie v analiz rechi. Razlichitel'nye priznaki i ikh korrelyaty. Glava 2. Opyt opisaniya razlichitel'nykh priznakov [Preliminaries to speech and its correlates. Chapter 2. Experience of describing distinctive features]

- analysis: the distinctive features and their correlates. Experience of describing distinctive characteristics. Ch. 2]. In: *Novoe v lingvistike* [New horizons in linguistics], Vol. 2. Moscow, pp. 173-230.
17. Eker S. (2007) Toplumdilbilgisel olarak /e/. In: *Edebiyat ve Dil Yazıları. Mustafa İsen'e Armağan*. Ankara, pp. 231-247.
 18. Fazli C., Patton Jon M. (2010) Change of word characteristics in 20-century Turkish literature: a statistical analysis. *Journal of Quantitative Linguistics*, 17 (3), pp. 167-190.
 19. Greenberg J.H. (1960) A quantitative approach to the morphological typology of language. *International Journal of American Linguistics*, 26, pp. 178-194.
 20. Trnka B. A (1966) *Phonological analysis of present day standart English*. Tokyo: Hokuon Publishing Go.
 21. Twaddell W.P. (1938) A phonological analysis of intervocalic consonant clusters in modern German. *Actes du IV Congrès international de linguistes à Copenhague*, pp. 218-225.
 22. Zolkhoev V.I. (2006) *Klassifikatsiya fonem posredstvom tsifrovogo koda: v 12 t.* [The classification of phonemes: digital code application: in 12 vol.]. Ulan-Ude: Buryat State University.
 23. Zipf G.K. (1932) *Selected studies of the principle of relative frequency in language*. Cambridge, Mass.: Harvard University Press.
 24. Zubkova L.G. (1990) *Fonologicheskaya tipologiya slova* [Phonological typology of the word]. Moscow: Peoples' Friendship University.