

УДК 811.512.14

ГРАММАТИЧЕСКАЯ И СТАТИСТИЧЕСКАЯ СТРУКТУРА ТЕКСТА НА АГГЛЮТИНАТИВНОМ ЯЗЫКЕ В ИСТОРИЧЕСКОЙ ПЕРСПЕКТИВЕ (на материале староосманских и турецких текстов)

Вадим Борисович Касевич

д. филол. н., почетный профессор СПбГУ

зав. кафедрой междисциплинарных исследований в области языков и литературы,
зав. кафедрой теории и методики преподавания языков и культур Азии и Африки

Санкт-Петербургский государственный университет

190000, Санкт-Петербург, Университетская наб., д. 11. vadim.kasevich@gmail.com

Аполлиария Сергеевна Аврутинак. филол. н., доцент кафедры междисциплинарных
исследований в области языков и литературы,

доцент кафедры теории и методики преподавания языков и культур Азии и Африки

Санкт-Петербургский государственный университет

190000, Санкт-Петербург, Университетская наб., д. 11. apollinaria1@yandex.ru

Настоящая статья посвящена выявлению количественно-типологических характеристик текста на агглютинативном языке в исторической перспективе. Сопоставляются соответствующие индексы, вычисленные по тексту на османском языке (турецкий язык до начала XX в.) и по тексту на турецком, что позволит установить количественные характеристики изменений, свойственных данному агглютинативному языку в диахронии. Дается динамический анализ агглютинативных языков, представленных материалом османского языка и современного турецкого. Текст на указанных языках рассматривается с точки зрения информационных показателей и формальной структуры. Авторы статьи подсчитали специальные количественные индексы для каждого текста, что делает возможным формализовать релевантные признаки рассматриваемых языков. Сравнение индексов позволит понять, каким образом происходят структурные изменения в исторически связанных между собой агглютинативных языках с точки зрения типологических характеристик текстов, анализируемых с позиций исторической перспективы. Настоящая статья является одной из серии публикаций в рамках проекта, посвященного развитию турецкого языка в исторической перспективе.

Ключевые слова: турецкий язык; османский язык; тюркская лингвистика; агглютинативные языки; статистическая структура; метод индексальной типологии.

Из истории вопроса, некоторые теоретические аспекты индексальной типологии, постановка проблемы

Историки языкознания помнят, что свою вступительную лекцию в Императорском Санкт-Петербургском университете, прочитанную 21 сентября 1900 г., И. А. Бодуэн де Куртенэ озаглавил «О смешанном характере всех языков» [Бодуэн де Куртенэ 1963]. Имелись в виду ареальный и исторический аспекты вопроса: Бодуэн де Куртенэ утверждал, что любой язык в процессе своей многовековой эволюции претерпевает влияние со стороны других языков, воспринима-

ет те или иные их черты и, как следствие, приобретает «смешанный» характер.

С суждениями Бодуэна де Куртенэ перекликаются высказывания Э. Сепира (1993), но только уже в плане типологических классификаций: Сепир подчеркивает тот известный факт, что, строго говоря, не существует языков чисто флективных, чисто агглютинативных, чисто изолирующих – во всяком языке можно найти элементы флективности, агглютинативности, изоляции; даже в рамках конкретной парадигмы могут сосуществовать типологически разные словоформы. Оставляя в стороне иллюстрации Гринберга,

можно вспомнить, что в урду образование множественного числа существительных носит в целом агглютинативный характер, но в арабских заимствованиях сохраняется так называемое ломаное множественное языка-донора¹. В русском языке во временной парадигме несовершенного вида глаголы настоящего и прошедшего времени носят синтетический характер (*читаю, читал*), а глаголы будущего времени – аналитический (*буду читать*). Различается, однако, **соотношение** типологически разнородных элементов в языках: в одном преобладает флексия и меньше представлена агглютинация, в другом – наоборот.

Коль скоро речь идет о **преобладании**, естественно должна возникнуть мысль о **количественном выражении**, о мере такого преобладания. Именно к этой мысли пришел Дж. Гринберг в своей известной работе 1960 г. [Greenberg 1960] (русский перевод 1963 г., см.: [Гринберг 1963]). Ученый предложил оценивать тип языка с помощью квантитативно-типологических индексов, каждый из которых представляет собой дробь, числителем которой выступает число, например, префиксов в некотором репрезентативном 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 языков².

Чрезвычайно важно, что сравнимость индексов и, соответственно, языков должна обеспечиваться достаточно строгими универсальными определениями тех единиц, в терминах которых выражаются типологические индексы. При невыполнении указанного условия всякая типология, квантитативная в особенности, теряет смысл. Описываемая методика была усовершенствована в монографии [Квантитативная типология языков... 1982]. Метод, предложенный Дж. Гринбергом и известный в литературе как квантитативная типология, может быть также назван *методом индексальной типологии*.

По своей логической природе предварительно описанные выше типологические индексы могут отличаться друг от друга. Так, Гринберг при оценке степени проявленности агглютинации в данном тексте (и, соответственно, языке) исходит не из числа агглютинативных аффиксов, а из числа агглютинативных морфемных (морфных) **швов**. Для него индекс агглютинации имеет вид A/J, т.е. отношение числа агглютинативных швов (число определяется тривиально по формуле $n - 1$, где n – число слов) к общему числу швов. В отличие от него, уже упоминавшийся индекс префиксальности носит иную природу.

Сложив агглютинативные и неагглютинативные швы, мы должны получить общее число швов в тексте, однако число префиксов «не с чем» складывать. Причина различий понятна: если морфемный шов появляется автоматически при любом сочетании морфем³, то этого же никак нельзя сказать о префиксах; префикс не является обязательным компонентом словоформы.

Если набор индексов является оптимальным, а текст – репрезентативным, то результат квантитативно-типологического исследования реализуется как установление множества параметров-индексов⁴, дающих комплексную многостороннюю типологическую характеристику данного языка с учетом «веса» каждого из учитываемых явлений в общей структуре оцениваемого языка.

Разумеется, величины индексов могут варьировать не только на границах классов более или менее однородных языков, но и внутри таких классов. Однако, как показал материал 1982 г. [Квантитативная типология языков... 1982], сравнительно легко обнаруживаются резкие перепады значений, отвечающие границам между классами.

Иначе можно сказать, что выделяемые по квантитативным основаниям классы носят не только количественный, но и *категориальный* характер. Напрашивается аналогия с категориальным восприятием фонем. Еще в экспериментах 1960-х гг. установлено, что если предъявлять испытуемым акустические стимулы, монотонно изменяющиеся по некоторому параметру (соответствующему, например, значению второй форманты), то в каком-то интервале стимулы воспринимаются как идентичные; однако при этом обычно обнаруживается точка равновероятного отнесения стимула к парадигматически «соседним» фонемам, которая и отвечает противоположенности в системе соответствующих фонем как категорий⁵.

Как следует из названия статьи, нас интересуют преимущественно языки, которые традиция относит к *агглютинативным*. В отличие от преобладающей традиции, мы делим множество аффиксов не на два: агглютинативные / флективные, а на **четыре** подмножества (подкласса): агглютинативные / неагглютинативные / флективные / нефлективные.

Агглютинативный аффикс удовлетворяет условию: «одно грамматическое значение – один аффикс»; в противном случае аффикс квалифицируется как неагглютинативный. Иначе говоря, чтобы ответить на вопрос, является ли данный аффикс агглютинативным, нужно выяснить, существуют ли в рассматриваемом языке другие аффиксы, которые передавали бы

то же значение – по существу, выступали бы как синонимы по отношению к данному. Из положительного ответа на сформулированный вопрос следует, что рассматриваемый аффикс является агглютинативным, из отрицательного, соответственно – неагглютинативным. Например, английская форма прошедшего времени с показателем *-ed* должна квалифицироваться как агглютинативная. Несмотря на формы типа *took, ate, slept* и т.п., последние выделяются как исключения, «неправильные» формы; «по умолчанию» (например, при формировании словоформ прошедшего времени от инновативных глаголов) выбираются варианты с *-ed*⁶. В отличие от этого, русские окончания множ. числа *-Ø (нож), -а (рука)* и т.п. выступают неагглютинативными: как показывают простые примеры выше, множественное число существительных русского языка обслуживается окончаниями-синонимами.

Флективный аффикс выделяется тем, что он вводит соответствующую словоформу не менее чем в две оппозиции. Например, *-а* в словоформе *рука* вводит эту словоформу в две оппозиции: по падежу (*рука – руку* и т.д.) и по числу (*рука – руки*). Соответственно, нефлективный аффикс «отвечает» за одну оппозицию. Например, префикс *вс-* в словоформе *вспахать* обеспечивает лишь одно противопоставление – видовое.

В результате мы получаем четыре признака классификации аффиксов: агглютинативность, неагглютинативность, флективность, нефлективность. Поскольку классы аффиксов выделяются по разным основаниям, принадлежность конкретного аффикса к одному из классов не исключает его одновременного вхождения в другой. С учетом незначимости порядка применения признаков в итоге мы получаем следующую систему аффиксов:

1) агглютинативные, нефлективные; 2) агглютинативные, флективные; 3) нефлективные, неагглютинативные; 4) флективные, неагглютинативные.

Каждому признаку отвечает особый типологический индекс: A/AF , $nonA/AF$ и т. д., где A – число агглютинативных аффиксов, AF – общее число аффиксов и т. п. Понятно, что при таком подходе агглютинативный язык – это язык с достаточно высоким индексом A/AF и достаточно низким значением индекса F/AF .

Остается вычислить эти индексы. В результате мы получим то, что можно обозначить как пересечение грамматической и статистической структур соответствующих текстов: индексы построены на специальных понятиях, важных для грамматики в типологическом аспекте, но одно-

временно они отражают распределение соответствующих параметров в пространстве текста.

Квантитативный (индексальный) подход позволяет распространить эту методику на **диахронические исследования**. Связь исторического и типологического языкознания осознана не так давно (см., например: [Оглоблин 2009]). Между тем она очевидна. В частности, эта связь проявляется (*может* проявляться) в том, что, изменяясь в процессе эволюции, язык в одних случаях сохраняет свою типологическую «физиономию», в других же изменяет ее – переходит из одного типологического класса в другой. Учитывая известную зыбкость границ между разными типологическими классами, о которой говорилось выше, именно введение своего рода типологической «метрики» обеспечит объективную базу для историко-типологических штудий⁷.

Множество индексов и множество текстов. Формирование диахронически ориентированного множества типологических индексов

Выбор типологических индексов, историческая динамика которых отражала бы наиболее естественным образом типологические тренды языковой системы, задача не тривиальная. Как можно было видеть из вышесказанного, какой бы то ни было традиции, которой можно было бы следовать, здесь просто нет. Мы опирались на результаты собственных исследований [Грамматика и семантика... 2011], а также на некоторые общие соображения и естественные следствия из теоретических представлений, широко распространенных в классическом языкознании.

В конечном итоге мы сформировали некоторое множество индексов, предположительно отражающих историко-типологическую динамику. Ниже представлен соответствующий перечень с необходимыми краткими комментариями.

Первый индекс, который мы хотели бы здесь ввести, – это

(1) **Индекс асиллабичности**. Индекс вычисляется как Syl_0/M , т.е. дает информацию об удельном весе неслоговых (асиллабических) морфем в соответствующем языке. Типологический смысл данного индекса заключается прежде всего в том, что отсутствие асиллабических морфем типа рус. *-л, -к-*, англ. *-z* и т.п. – важный признак слоговых (силлабических, силлабемных) языков [см.: Касевич 1983].

(2) **Индекс полисиллабичности**. Он вычисляется как Syl_n/W , т.е. отношение n -сложных слов к общему числу слов в тексте. Иначе говоря, отдельно подсчитывается процент (удельный вес) односложных, двусложных, трехсложных,

четырёхсложных слов. Этот же индекс известен как длина слова [Москович 1967].

Смысл введения данного индекса заключается в том, что из некоторых известных реконструкций (см. например: [Щербак 1981]) как будто бы следует, что большинство слов пратюрского языка были односложными, что характеризует этот язык вполне определенным образом – причем не только фонологически и морфонологически, но и морфологически (подробнее об этом см. ниже). В литературе описаны и противоположные по направлению процессы. Так, континентальные австронезийские языки, попав в ареал слоговых (силлабемных), утратили типичную для них двусложную структуру морфемы, заменив ее, как результат стяжения двуслога, на односложную, ср., например: *gugu* → *ku* (см.: [Касевич 1983]).

(3) **Индекс глубины слова.** Данный индекс определяется как M/W , т.е. как отношение числа морфем к числу слов (ср. индекс синтеза у Дж. Гринберга, а также: [Москович 1967]).

(4) **Индекс агглютинативности.** Об этом индексе достаточно говорилось выше. Толкование агглютинативности, изложенное ранее, предполагает, что с повышением индекса агглютинативности снижается использование синонимических аффиксов и повышается, соответственно, представленность безвариантных или алломорфологически варьирующих аффиксов. Роль этого показателя для определения типологического «покроя» языка не требует пояснений.

(5) **Индекс флективности.** Значение индекса (5) уже объяснялось и не будет здесь воспроизводиться. Чрезвычайно существенно, что все введенные здесь индексы оказываются логически связанными друг с другом. Действительно, если каждая служебная морфема передает ровно одно значение (отвлекаясь от полисемичности), то «прибавление» в рамках слова каждого дополнительного значения предполагает добавление морфемы. Если такая служебная морфема является одно- и более сложной, то увеличивается как глубина, так и длина слова. Если морфема носит асиллабический характер, то возрастает только глубина слова при его сохранной длине.

(6) **Индекс сохранения базового порядка слов.** В отличие от предыдущих, данный индекс ориентирован на синтаксические характеристики языка. Реально имеются в виду два «подынка», которые, возможно, в дальнейшем будут трактоваться отдельно. Первый из них – это базовая синтаксическая структура предложения: SOV или SVO; индекс здесь принимает вид $ORDN_{nat}/SENT$, т.е. отношение предложений с «естественным» порядком слов к общему числу

предложений. Структуры SOV, SOV, VSO отражают исторически и жанрово устойчивый параметр. Тем не менее известно, что степень его устойчивости в разных языках и жанрах может различаться. Если, например, для бирманского языка соблюдение базовой структуры SOV есть непреложный закон, то в современной турецкой поэзии вынесение глагола-сказуемого в начальную позицию отнюдь не исключается. Второй синтаксический «подынок» – это тип маркирования синтаксических отношений в именных сочетаниях типа NN. По данным Дж. Никольс [Никольс 2006] тип такого маркирования оказывается типологически релевантным для самых разных языков; зная этот тип, мы можем вывести целый ряд типологических свойств языка, которые «на поверхности» как будто бы никак не связаны с данным типом и друг с другом. Формальное выражение этого индекса имеет вид NN/SENT (с необходимыми диакритиками, уточняющими тип маркирования).

(7) **Индекс исконности.** Значение этого индекса вычисляется как Frn/W , т.е. как отношение заимствованных слов к общему числу слов в тексте. Здесь мы покидаем сферу грамматики и переходим в область параметров **словаря**. Строго говоря, этот индекс не является формально-типологическим, однако передаваемая им характеристика вряд ли будет излишней.

Результаты работы по вычислению типологических индексов – это получение множества значений индексов и установление разницы между ними.

Языковой материал для апробации методики индексальной типологии в диахронии

Термин «языковой материал» используется здесь в смысле Щербы: как коллекция текстов, анализ которых позволяет установить систему языка или ее фрагмент. Мы выбрали для такого рода анализа два параллельных текста: один на староосманском языке и другой на современном турецком. Это роман «Приключения Али-бея, или Предосторожность» («Sergüzeşt-i Ali Bey, veya İntibah»), принадлежащий Намык Кемалю, по мнению большинства исследователей, первый роман в истории турецкой литературы (впервые опубликован в 1876 г., [Namik Kemal 1876]) [Kabaklı 2008]. Современный текст представляет собой перевод староосманского [Namik Kemal 2012], временная дистанция между ними – примерно 100 лет. Представляется, что сравнение двух такого рода генетически связанных идиом оправданно. Хотя указанного промежутка времени вряд ли достаточно, чтобы между ними (идиомами) появились типологически зна-

**Касевич В. Б., Аврутина А. С. ГРАММАТИЧЕСКАЯ И СТАТИСТИЧЕСКАЯ СТРУКТУРА
ТЕКСТА НА АГГЛЮТИНАТИВНОМ ЯЗЫКЕ В ИСТОРИЧЕСКОЙ ПЕРСПЕКТИВЕ**
(на материале староосманских и турецких текстов)

чимые, тем более яркие различия, индексальная типология может оказаться более тонким инструментом для обнаружения тенденций типологического развития языков. Кроме того, существует мнение о том, что тюркские языки изменяются – по сравнению с языками «иной генетики» – сравнительно быстро и в известной степени противостоят таким языкам, как исландский или литовский; последний, по известному суждению Соссюра, за 500 лет практически не изменился.

Термин «турецкий язык» в современной специальной литературе означает современный государственный язык Республики Турция, которая была провозглашена в 1923 г. в результате распада Османской империи после ее участия в Первой мировой войне и победы Национально-освободительного движения.

До XX в. в Османской империи существовал литературный язык – османский, главной особенностью которого было обилие (до 90 %) лексических и нередко синтаксических арабо-персидских заимствований. Этот литературный язык начал складываться на рубеже XV–XVI вв. на основе староанатолийско-тюркского языка (языка огузо-сельджукских тюркских племен, которые жили в Средней Азии, но в X в. были вытеснены оттуда иными тюркскими племенами – уйгурами). Староанатолийско-тюркский язык сформировался к XI–XII вв., когда огузские племена из Средней Азии окончательно переселились в Малую Азию; в языковом плане новый язык соединил язык огузов Средней Азии и язык смешанного тюркского населения Анатолии.

Литературный османский язык оказал существенное влияние и на повседневный разговорный язык – *тюркю*, в котором также велико было число заимствований, обильно использовались различные арабские и персидские синтаксические конструкции, чуждые структуре тюркских языков.

После основания Республики Турция в 1923 г. начался процесс очистки турецкого языка от ар-

хаичных арабских и персидских слов и замены их исконными турецкими словами, которые нередко создавались лингвистами на базе материала древних тюркских языков. Специально для этого в 1932 г. было создано государственное «Турецкое лингвистическое общество», главной целью которого были и остаются тюркизация и модернизация турецкого языка. Процесс «очищения» турецкого языка продолжается и в наши дни.

Этот процесс привел к стремительному изменению лексического состава языка. В результате представители разных поколений подчас с трудом понимают друг друга; более того, молодежь, которая пользуется современным турецким языком, с трудом читает книги, написанные даже в 1930–1940-х гг., не говоря уже о книгах, написанных в XIX в.⁸ Например, нынешние издания турецкого романиста Сабахаттина Али, писавшего в 1930-е и 1940-е гг., сопровождаются обильными сносками, представляющими перевод устаревших слов на их новые эквиваленты.

Сравнительный анализ индексов текстов на староосманском и турецком языках

Итак, рассмотрим числовые характеристики для обоих текстов, а затем сопоставим их.

Для сравнительного анализа был выбран один и тот же фрагмент романа. Приблизительный размер фрагмента в обоих текстах – 500 слов, по смыслу они идентичны, однако по размеру несколько различаются (примерно на 30 слов «в пользу» староосманского), т. к. в архаичном тексте чаще употребляются перифрастические конструкции с иноязычным смысловым существительным и вспомогательным глаголом в личной финитной форме.

Текст на староосманском языке далее фигурирует как «текст 1»; текст на современном турецком – как «текст 2».

Данные по индексам сведены в таблицу.

Типологические индексы для текста 1 и текста 2

Наименование индекса	Староосманский (текст 1)	Современный турецкий (текст 2)
<i>Индекс асиллабичности</i>	0	0
<i>Индекс моносиллабичности</i>	0.18	0.18
<i>Индекс дисиллабичности</i> (двусложности)	0.4	0.3
<i>Индекс глубины слова</i>	1.54	1.8
<i>Индекс агглютинативности</i>	1	1
<i>Индекс флективности</i>	0	0
<i>Индекс исконности</i>	0.48	0.22

Поясним и прокомментируем данные, представленные в таблице.

Выделим значения индексов, свидетельствующие о максимальной устойчивости лексико-грамматических характеристик, которую обнаруживают во времени рассматриваемые идиомы. Прежде всего, это относится к базовым типологическим параметрам, связанным с агглютинативностью / флективностью. При статистическом обследовании текстов мы не обращались к расщеплению признака «агглютинативность / флективность», о целесообразности чего говорилось в начале статьи. Причина ясна: в обоих текстах полностью отсутствуют как неагглютинативные, так и флективные аффиксы, не существуют (во всяком случае в пределах текстов 1 и 2) ни синонимия аффиксов, ни их «бифункциональность».

Столь же высокую устойчивость обнаруживают индексы, связанные с морфонологическими, в сущности, параметрами – отсутствием асиллабических и постоянством числа моносиллабических слов. В тексте 2 несколько ниже значение индекса дисиллабичности при сохранении индекса моносиллабичности; очевидно, это объясняется снижением числа заимствований, для которых в целом характерна относительная многосложность.

У нас нет гипотезы, которая могла бы убедительно объяснить некоторое повышение индекса глубины слова при переходе от староосманского к современному турецкому (см. таблицу). Возможно, объяснение заключается в том, что в современном турецком больше используются грамматические показатели – меньше «работает» известный принцип факультативности грамматических элементов.

Как можно видеть, мы не использовали в обработке данных индекс сохранения базового порядка слов. Данный индекс оказался для наших текстов нерелевантным: в обоих текстах нарушение базового порядка слов наблюдалось в ничтожном числе случаев.

Не вошел в таблицу стоящий особняком индекс, который можно было бы назвать «индексом атрибутивности»; в числителе у этого индекса – сумма числа прилагательных и изафетных конструкций, в знаменателе – число предложений. Для текста 1 индекс атрибутивности равен 1.3, для текста 2 – 0.8. По-видимому, за этим различием стоят стилистические факторы: автор архаического текста часто стремится уснащать существительные разного рода определениями, в то время как «модерну» (в широком смысле) свойствен известный минимализм.

Наконец, самое большое различие (предсказуемое) между текстом 1 и текстом 2 проявляется в значениях индекса исконности: для староосманского языка этот индекс равен 0.48, для современного турецкого – 0.22, т.е. более чем в два раза меньше.

Разумеется, можно сказать, что принципиально те же результаты можно было бы получить и без применения сравнительно громоздкого аппарата индексальной типологии. Кроме того, полученные данные не были обработаны с точки зрения статистических критериев оценки значимости различий между средними и иными величинами. Но надо заметить, что метод был применен для оценки соотношения весьма близких во всех отношениях идиом, и нашей задачей было испытание этого метода в своего рода граничной ситуации. Представляется, что, дополненный математико-статистической обработкой данных, этот метод может стать достаточно надежным и тонким способом оценки соотношения идиом в компаративных исследованиях.

Примечания

¹ Согласно Дж. Гринбергу, который агглютинацию определял по отношению «одно грамматическое значение – одно формальное средство выражения», уже сам по себе факт наличия двух средств передачи множественности (или иной семантики) говорит о неагглютинативности категории числа как таковой. Однако при таком подходе ликвидируется понятие *исключения*, что вряд ли оправданно [см. об этом: Гринберг 1963; Квантитативная типология языков... 1982; Касевич 1988].

² Понятно, что на численные величины влияли не только типологические параметры соответствующих языков, но и параметры текстов, подобранных для исследования. Это отдельная проблема, которой касаться не будем.

³ Здесь мы допускаем определенное упрощение: более сложна картина в тех языках, где существуют разрывные аффиксы и / или разрывные корни (как в семитских языках).

⁴ В принципе индексы должны выступать как независимые показатели, но определение их независимости – это особая формальная и содержательная проблема, которая здесь не исследуется (см. об этом: [Яхонтов К. С., Яхонтов С. Е. 1986]).

⁵ Как было позднее обнаружено, элементы категориального восприятия имеются уже у животных (см. об этом, например: [Фитч 2013]).

⁶ Эксперименты Дервинга и Скоусена показали, что при наличии выбора между регулярной («правильной») формой и иррегулярной («непра-

вильной») испытуемые всегда выбирают регулярную [Derwing, Skousen 1989].

⁷Здесь налицо и своего рода психологическая составляющая: наблюдая язык в его исторической динамике, исследователь не может не отмечать материальные изменения, но порой «не видит» структурные сдвиги, изменяющие «покрой» языка.

⁸См. об этом: [Аврутина, Глазанова] (в печати).

Список литературы

Аврутина А. С., Глазанова Е. В. О количественной оценке сравнительной трудности восприятия разновременных текстов (в печати).

Бодуэн де Куртенэ И. А. О смешанном характере всех языков // Бодуэн де Куртенэ И. А. Избранные труды по общему языкознанию. М.: АН СССР, 1963. Т. 1. С. 362–372.

Грамматика и семантика восточного текста: Квантитативные характеристики / под ред. В. Б. Касевича. СПб.: Изд-во РХГА, 2011. 214 с.

Гринберг Дж. Квантитативный подход к морфологической классификации языков // Новое в лингвистике. М.: Изд-во иностр. лит. 1963. Вып. III. С. 60–94.

Касевич В. Б. Семантика. Синтаксис. Морфология. М.: Наука, 1988. 309 с.

Касевич В. Б. Фонологические проблемы общего и восточного языкознания. М.: Наука, 1983. 295 с.

Квантитативная типология языков Азии и Африки / под ред. В. Б. Касевича, С. Е. Яхонтова Л.: ЛГУ, 1982. 331 с.

Москович В. А. Глубина и длина слова в естественных языках // Вопросы языкознания. 1967. №6. С. 17–33.

Никольс Дж. Вершинное / зависимостное кодирование / под ред. А. А. Кибрика, И. М. Кобозевой, И. А. Секериной // Современная американская лингвистика: Фундаментальные направления. М.: Едиториал УРСС, 2006. С. 293–301.

Оглоблин А. К. Очерк диахронической типологии малайско-яванских языков. М.: УРСС, 2009. 200 с.

Сепир Э. Избранные труды по языкознанию и культурологии. М.: Прогресс – Универс, 1993. 656 с.

Фитч У. Т. Эволюция языка. М.: Языки славянской культуры, 2013. 768 с.

Щербак А. М. Очерки по сравнительной морфологии тюркских языков: Глагол. Л.: Наука, 1981. 92 с.

Яхонтов К. С., Яхонтов С. Е. Применение корреляционного и факторного анализа к типологии языков // Актуальные вопросы дериватологии

и дериватологии / под ред. Б. И. Барткова. Владивосток: ДВО АН СССР, 1986. С. 105–125.

Derwing B., Skousen R. Morphology in the mental lexicon: A new look at analogy // Yearbook in Morphology: Journal of Psycholinguistic Research. 2. Amsterdam, 1989. P. 55–71.

Greenberg J. H. A quantitative approach to the morphological typology of languages // International Journal of American linguistics. 1960. Vol. 26, №3. P. 178–194.

Kabaklı A. Türk Edebiyatı. 3. cilt. İstanbul: Türk Edebiyatı Vakfı Yayınları, 2008. 540 s.

Namık Kemal. İntibah, Sergüzeşt-i Ali Beyi, İstanbul, 1876. 222 s.

Namık Kemal. İntibah. İstanbul: İskele, 2012. 224 s.

References

Avrutina A. S., Glazanova E. V. O kolichestvennoj ocenke sravnitel'noj trudnosti vosprijatija raznovremennykh tekstov [On the quantification of the relative difficulty of multi-temporal perception texts] (forthcoming).

Boduen de Kurtene I. A. O smeshannom kharaktere vseh jazykov [On the mixed nature of all languages]. Boduen de Kurtene I. A. Izbrannye trudy po obshhemu jazykoznaniju. [Selected works on general linguistics] Vol. 1. Moscow, USSR Academy of Science Publ., 1963. P. 362–372.

Derwing B., Skousen R. Morphology in the mental lexicon: A new look at analogy. Yearbook in Morphology. Journal of Psycholinguistic Research. 2. Amsterdam, 1989. P. 55–71.

Fitch U. T. Eholvojucija jazyka [The Evolution of Language]. Moscow, Jazyki slavjanskoj kul'tury Publ, 2013. 768 p.

Grammatika i semantika vostochnogo teksta: Kvantitativnye kharakteristiki. [Grammar and semantics of the eastern Text: Quantitative characteristics]. Ed. by V. B. Kasevich. St. Petersburg: Christian Russian State Academy Publ., 2011. 214 p.

Greenberg J. H. A quantitative approach to the morphological typology of languages. International Journal of American linguistics. 1960. Vol. 26, no 3. P. 178–194.

Grinberg G. Kvantitativnyj podkhod k morfoloicheskoj klassifikacii jazykov [Quantitative approach to the morphological classification of languages]. Novoe v lingvistike. [New in linguistics] Iss. III. Moscow, Inostrannaja literatura Publ., 1963. P. 60–94.

Jakhontov K. S., Jakhontov S. E. Primenenie korrelacionnogo i faktornogo analiza k tipologii jazykov [The use of correlation and factor analysis to language typology]. Aktual'nye voprosy deriva-

tologii i derivatografii [Topical questions and word formation derivatography]. Ed. by B. I. Bartkov. Vladivostok, 1986. P. 105–125.

Kabaklı A. Türk Edebiyatı. 3. cilt. İstanbul: Türk Edebiyatı Vakfı Yayınları, 2008. 540 s.

Kasevich V. B. Fonologicheskie problemy obshhego i vostochnogo jazykoznanija [Phonological problems of general and Eastern linguistics]. Moscow, Nauka Publ., 1983. 295 p.

Kasevich V. B. Semantika. Sintaksis. Morfologija [Semantics. Syntax. Morphology]. Moscow, Nauka Publ., 1988. 309 p.

Kvantitativnaja tipologija jazykov Azii i Afriki [Quantitative typology of Asian and African languages] / Ed. by V. B. Kasevich, S. E. Jahontov. Leningrad, LSU Publ., 1982. 331 p.

Moskovich V. A. Glubina i dlina slova v estestvennykh jazykakh [The depth and length of words in natural languages]. Voprosy jazykoznanija. [Questions of Linguistics] 1967. No 6. P. 17–33.

Namik Kemal. İntibah, Sergüzeşt-i Ali Beyi, İstanbul, 1876. 222 s.

Namik Kemal. İntibah. İstanbul: İskele, 2012. 224 s.

Nikol's Dzh. Vershinnoe / zavisimostnoe kodirovanie. [Top / connected coding]. Ed. by A. A. Kibrik, I. M. Kobozeva, I. A. Sekerina. Sovremennaja amerikanskaja lingvistika: Fundamental'nye napravlenija [Contemporary American linguistics: Fundamental Directions] Moscow, Editorial URSS Publ., 2006. P. 293–301.

Ogloblin A. K. Oчерk diakhronicheskoj tipologii malajsko-javanskikh jazykov. [Essay of diachronic typology in Malay-Javanese]. Moscow, URSS Publ., 2009. 200 p.

Sepir E. Izbrannye trudy po jazykoznaniju i kul'turologii [Selected Papers on linguistics and cultural studies]. Moscow, Progress – Univers Publ., 1993. 656 p.

Shherbak A. M. Oчерki po sravnitel'noj morfologii tjurkskikh jazykov: Glagol. [Essays on comparative morphology of the Turkic languages: Verb]. Leningrad, Nauka Publ., 1981. 92 p.

**GRAMMATICAL AND STATISTICAL STRUCTURE OF A TEXT
IN THE AGGLUTINATIVE LANGUAGE IN HISTORICAL PERSPECTIVE
(basing on Turkish and Ottoman texts)**

Vadim B. Kasevich

**Professor, Professor Emeritus of Saint Petersburg State University,
Head of the Department of Interdisciplinary Studies in the Field of Languages and Literature,
Head of the Department of Theory and Methods of Teaching Languages
and Cultures of Asian and African Countries
Saint Petersburg State University**

Apollinaria S. Avrutina

**Associate Professor in the Department of Interdisciplinary Studies in the Field of Languages and Literature,
Associate Professor in the Department of Theory and Methods of Teaching Languages
and Cultures of Asian and African Countries
Saint Petersburg State University**

This article focuses on identification of quantitative typological characteristics of a text in the agglutinative language in historical perspective. The article compares the corresponding indices calculated from a text in the Ottoman language (the Turkish language before 1928) and a text in modern Turkish, which will make it possible to establish quantitative characteristics of changes attributable to the agglutinative languages in diachrony.

The paper presents a dynamic diachronic analysis of the agglutinative language. Both texts are analyzed in terms of information indicators and formal structure. The authors calculated specific quantitative indices for each of the texts, which makes it possible to formalize relevant features of the languages in diachrony. Comparison of indices allows for understanding of how structural changes occur in historically linked agglutinative languages in terms of typological characteristics of texts in historical perspective. This article is one of a series of publications in the framework of a project on evolution of the Turkish language in historical perspective.

Key words: Turkish language; Ottoman language; Turkic linguistics; agglutinative languages; statistical structure; index typology method.