



АРАБИСТИКА ЕВРАЗИИ
EURASIAN ARABIC STUDIES
ЕВРАЗИЯ АРАБИСТИКАСЫ

2020, No. 12

ISSN 2619-1261

УДК 811.411.21

Оригинальная Статья
Original Paper

АНАЛИЗ ЯЗЫКА КОРАНА НА ОСНОВЕ АВТОМАТИЧЕСКИХ МЕТОДОВ ОБРАБОТКИ ДАННЫХ

О.И. Редькин¹, О.А. Берникова²

Санкт-Петербургский государственный университет

o.redkin@spbu.ru¹, o.bernikova@spbu.ru²

Поступила в редакцию: 15 сентября 2020 г. Submitted: September 15, 2020
Одобрена рецензентами: 18 октября 2020 г. Reviewed: October 18, 2020
Принята к публикации: 22 октября 2020 г. Accepted: October 22, 2020

Аннотация

Изучение языка Корана привлекало внимание исследователей, начиная с работ первых арабских грамматистов, которые рассматривали особенности его лингвистической структуры, в том числе и количественные параметры. Появление в последние десятилетия компьютерных технологий и их применение в области текстологического анализа открыло новую страницу в истории изучения Корана. Вместе с тем имеющиеся решения, направленные на детальное изучение его отдельных характеристик на различных лингвистических уровнях не в полной мере отвечают требованиям, предъявляемым к процессингу арабографических текстов. В данном исследовании на примере глаголов десятой породы был проведен инновационный анализ особенностей языка Корана. На базе разработанных для этого технологий с использованием стандартного программного обеспечения решались задачи, связанные с автоматизированным выделением глагольных словоформ в десятой породе, их классификацией в зависимости от показателей рода, числа, лица, а также категорий залога, времени и склонения. Результаты исследования представлены в виде таблиц и снабжены примерами. Статистические данные формируют наглядную картину реализации рассмотренных моделей, что, с одной стороны, открывает новые возможности для теоретического изучения арабского языка в диахронии, а с другой – позволяют применить апробированные решения

автоматической обработки арабографического контента на материале иных глагольных и именных словоформ.

Ключевые слова: *Коран, арабский язык, автоматическая обработка текста, формы глагола*

Для цитирования: *Редькин О.И., Берникова О.А. Анализ языка Корана на основе автоматических методов обработки данных // Арабистика Евразии. 2020. № 12. С.72-89.*

ANALYSIS OF THE LANGUAGE OF THE QUR'ĀN BASED ON AUTOMATIC DATA PROCESSING METHODS

O.I. Redkin¹, O.A. Bernikova²

Saint-Petersburg State University

o.redkin@spbu.ru¹, o.bernikova@spbu.ru²

Abstract

The study of the language of the Qur'ān has drawn the attention of researchers, starting with the works of the first Arab grammarians, who considered the features of its linguistic structure, including quantitative indicators. The emergence of computer technologies in recent decades and their application in the field of textual analysis has opened a new page in the history of the study of the Qur'ān. At the same time, the available solutions do not fully satisfy the requirements for the processing of the arabographic texts, aimed at a detailed study of individual characteristics of the sacred text at various linguistic levels. In this study, on the example of the X-th form of the verb, an innovative analysis of the features of the language of the Qur'ān was carried out. On the basis of the newly developed technologies, and with the implementation of standard software was solved a set of problems related to automated selection of the verbs of the tenth form, and their classification depending on the gender, number, person, as well as the categories of voice, time and mood. The research results are presented in the form of tables and provided with examples. The statistics provide a good picture of the implementation of the models examined, which, on the one hand, opens up new possibilities for theoretical study of the Arabic language in diachrony, and on the other hand, allow to apply proven solutions for the automatic processing of Arabic content on the material of other verb and nominal word forms.

Keywords: *Quran, Arabic language, automatic text processing, verb forms*

For citation: *Redkin, O.I., & Bernikova, O.A. (2020). Analiz yazyka Korana na osnove avtomaticheskikh metodov obrabotki dannykh [Analysis of the language of the Qur'ān based on automatic data processing methods]. Eurasian Arabic Studies, 12, 72-89. (In Russian)*

ВВЕДЕНИЕ

Появление первого списка Корана и дальнейшая его кодификация были вызваны необходимостью передачи ниспосланного Пророку откровения во времени и пространстве. Важность сохранения сакрального текста в неизменном виде стала особенно актуальной после включения в состав Халифата территорий за пределами Аравийского полуострова и принятия ислама населением, для которого арабский язык не был родным. Это, в свою очередь, стало импульсом к появлению первых арабских грамматических школ, а также созданию нормирующих грамматик, включающих набор правил и критериев, сформулированных на основе изучения эталонных текстов, в первую очередь, текста Корана.

Средневековые грамматисты рассматривали Коран как объект изучения, а также как своего рода нормативный текст, критерий языковых норм, включающий в себя всю палитру существующих грамматических правил, которые должны получить правильную трактовку и описание. В контексте изучения Корана уже в первых трудах средневековых авторов затрагивались вопросы, связанные с развитием графической системы арабского языка, синтаксиса и морфологии, лексики и фонетики. Так, в работах Абу ал-Асвада ад-Дуали (603-668), Абу Джа'фара ан-Наххаса (ум. 949), а также Абу Амра ибн ал-Ала ал-Басри (689-771), который являлся одним из основателей школ чтения Корана, Насра бин Асима ал-Ляйси (ум. в 708 г.) и автора самой ранней из дошедших до нашего времени грамматик – Абу Бишра Амра ибн Усмана ал-Басри (760-797), более известного как Сибавейхи, и сочинениях других представителей басрийской и куфийской лингвистических школ анализ Корана занимает центральное место.

В настоящее время изучение Корана, помимо его религиозной составляющей может дать информацию о морфологии, лексике и синтаксисе классического арабского языка, равно как и племенных диалектах и других семитских языках, существовавших на территории Аравии в конце 6 – начале 7 века н.э. Внимательное рассмотрение языкового контента позволяет получить и

косвенные данные об условиях, сопутствовавших формированию текста, в частности, целевой аудитории проповеди, ее эмотивной составляющей и хронологической соотнесенности отдельных аятов.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

История применения количественных методов анализа языка Корана восходит к работам первых арабских грамматистов, которые детально изучали не только содержательную сторону сакрального текста, но также и такие показатели, как количество аятов, слов и даже отдельных харфов. При этом следует отметить, что количественные оценки такого рода варьируются, что во многом зависит от того, рукописный или печатный вариант текста Корана используется для анализа, а также определяются индивидуальными критериями подсчета, которыми руководствуется исследователь.

Появление в 1923 году так называемого «каирского издания» Корана, которое кодифицировало графическую передачу сакрального текста, сделало возможным более точно определить количественные характеристики его отдельных компонентов.

Сегодня новые технологии позволяют по-новому подойти к решению ряда проблем, связанных с анализом текста Корана. Еще несколько лет назад создание программных продуктов, предназначенных для обработки больших объемов текста было компетенцией в первую очередь программистов, в свою очередь роль лингвистов заключалась в определении общих принципов обработки данных, актуальных для проведения такого анализа маркеров, а также проведении тестирования. Напротив, в настоящее время появляется все большее количество публикаций, отражающих результаты исследований, осуществленных при помощи программных продуктов, созданных самими лингвистами на основе наиболее распространенных языков программирования [Svensson, 2019, P. 211-229], а также с использованием имеющегося специализированного или стандартного программного обеспечения. Перечень такого рода работ включает исследования, посвященные проблемам лемматизации, частотности вхождений лексики, определения иерархии используемых в тексте морфологических моделей и др. При этом сфера исследований не ограничивается письменными текстами, созданными на основе кириллицы или латиницы, но и включает материалы, основанные на использовании других графических систем, в том числе и арабской графики.

В последние десятилетия совершенствование информационно-коммуникационных технологий дало новый толчок развитию гуманитарных исследований в целом. Использование объективных методов математического

анализа позволило успешно решать проблемы, связанные с изучением и интерпретацией текста Корана, его переводами на другие языки, особенностями его лексики, а также с рядом других задач.

В настоящий момент даже с помощью стандартного программного обеспечения можно оперативно решать некоторые задачи, ранее требовавшие значительных временных и материальных затрат, например, подсчет численности аятов, отдельных слов и даже букв.

Вместе с тем, решения, предлагаемые для текстов, основанных на латинице или кириллице, не всегда подходят для арабографического текста. Так, в ходе автоматической обработки материала на основе кириллицы или латиницы при вертикальной сегментации основными маркерами могут служить пробелы между словами и отдельными буквами.

В случае арабского языка данный процесс сегментации значительно сложнее, что обусловлено характером самой графической системы. Например, некоторые буквы имеют неслитное написание, что затрудняет сегментацию текста по критерию наличия / отсутствия пробелов. В силу этого, процессинг любого арабографического текста требует разработки специальных решений, учитывающих особенности арабского письма. Другими словами, графическая система арабского языка, характеризующаяся изменением написания букв в зависимости от их положения в линейной структуре слов, вариативностью оформления некоторых морфологических моделей, отсутствием в значительном числе случаев четких границ между служебными и значимыми словами, т.е. слитным написанием ряда предлогов и частиц, затрудняет сегментацию текста.

Помимо этого, анализ любого арабографического текста требует использования и учета всего объема имеющейся информации – как чисто лингвистического, так и экстралингвистического характера. Примером такого рода может служить процесс кодификации текста Корана и осуществленное в дальнейшем внесение диакритик и знаков для передачи кратких гласных, удвоения и т.д., постановка которых была основана на оценке возможных графических вариантов передачи словоформ, их соответствия правилам глагольной или именной морфологии, оформления в зависимости от синтаксической структуры текста, а также его семантики.

Важно отметить, что при создании единого для всей мусульманской уммы текста Корана помимо чисто лингвистической использовалась и информация экстралингвистического характера – свидетельства очевидцев, слышавших откровения пророка, обстановка в момент ниспослания отдельных сур и аятов и

др. Ведь «Коран – это запись живой речи, которая предполагала, что аудитория знает, о чем идет спор и способна сама дополнить рассказ необходимыми деталями» [Резван, 2001, С. 39].

Несмотря на то, что, будучи бесценным источником исторической, религиозно-ведческой и лингвистической информации на протяжении всей истории своего существования Коран всегда был объектом пристального внимания со стороны исследователей, анализ его текста с применением точных математических инструментов стал реальностью лишь сравнительно недавно. Последние десятилетия отмечены активным поиском в решении задач, связанных с дигитализацией арабографического текста и его интерпретацией, в том числе и текста Корана.

Так, значительным шагом на пути структурного описания лексики сакрального текста стало появление корпуса Корана [The Quranic Arabic Corpus], содержащего не только сведения относительно грамматического статуса каждого слова, но и предоставляющего информацию о количестве вхождений пород глаголов и их производных, а также позволяющего осуществлять поиск лексических единиц по корневой основе. Проект осуществлялся в Университете Лидса с 2009 по 2017 годы.

Использование корпуса Корана значительно расширяет сферу исследовательских задач, например, дает возможность получить информацию о наличии форм, соответствующих определенной породе, вместе с тем, данный ресурс не позволяет осуществить их подробную кластеризацию, т.е. детальную классификацию вхождений глаголов определенной породы в зависимости от их грамматических категорий.

Другая разработка Университета Лидса “Quran Analysis”, осуществленная под руководством Э. Атвела [Quran Analysis], позволяет проводить семантический анализ текста Корана с использованием интеллектуальных технологий обработки данных. Несмотря на применение ряда инновационных решений, данный ресурс не лишен ряда неточностей и недостатков, как показало его простейшее тестирование.

Важнейшим источником информации относительно лексики Корана и грамматических характеристик используемых словоформ является словарь “*Dictionary of Qur'anic Usage Online*” [Badawi & Abdel Haleem] издательства Brill Publishers. Работа, в которой подчеркивается роль контекста в определении значений каждой лексемы, основана на словарях классического арабского языка и комментариях к Корану. Разработке онлайн версии словаря предшествовало печатное издание «Арабско-английского словаря

контекстуального использования лексики Корана». Сами авторы, по их словам, «посвятившие свою жизнь кораническим штудиям», характеризуют его как «первый исчерпывающий, основанный на глубоком исследовании контекстных использований слов арабско-английский словарь Корана, а также составленный с учетом достижений современной лексикографии...» [Badawi & and Abdel Haleem]. Названный словарь дает не только примеры контекстуального использования коранической лексики, но и данные относительно количества вхождений, рассматриваемых именных и глагольных форм.

Одной из новейших разработок в области создания электронных ресурсов по коранистике стал учебно-методический комплекс «Лексико-грамматический строй Корана» А. Ю. Кудрявцевой и Е. А. Резвана. Ресурс размещён в открытом доступе и проходит стадию апробации [Кудрявцева, Резван].

Несмотря на наличие справочных материалов и других источников, содержащих информацию относительно количественных характеристик Корана, по-прежнему остается открытым вопрос о детальной классификации лексики в зависимости от формо- и словообразовательной модели.

В частности, речь идет об определении количества вхождений каждой из производных пород арабского глагола с учетом основных грамматических категорий (времени, наклонения, залога, а также лица, числа и рода).

Целью данного исследования является разработка технологии автоматической классификации арабских дериватов с использованием стандартного программного обеспечения. Объектом исследования выступают глаголы X породы, встречающиеся в тексте Корана.

Для проведения анализа использовался пакет приложений Microsoft Windows 10, с помощью которого осуществлялась разработка методики процессинга арабографического текста с целью классификации словоформ, производных от исходной морфологической модели и определения частотности вхождений каждой из них, а также анализа особенностей их реализации.

Созданный алгоритм представлял собой последовательность макросов, каждый из которых выполнял выделение соответствующих словоформ на основе определения маркеров глаголов-индикаторов грамматических категорий и на их основе проводил классификацию данных единиц.

Использование в качестве объекта исследования глаголов X породы обусловлено ее промежуточным по сравнению с глаголами других пород количеством вхождений. Данный выбор также связан с особенностями семантики данной породы, которая объединяет как значение возвратности, т.е. направленности на себя, так и каузатива, принудительности, направленности

действия на другой объект, что делает ее переходной. Имеют место также и другие оттенки значений данной породы, включая декларативность [Wright, 1896, P. 44-46].

Данная работа продолжает исследование, результаты которого были представлены в статье «Язык Корана: особенности морфологии глагола», опубликованной в 2019 г. [Берникова, Редькин, 2019, С. 59-72].

Как уже было отмечено ранее, процесс разработки алгоритмов (последовательности макросов) анализа графического текста Корана осложнялся самим характером арабской графики и трудностью ее формализации.

Так, часто отсутствовали графические границы между значимым и вспомогательным словом (слитное написание частиц и предлогов), имело место слитное написание постфиксов, варьирование гласных инфиксов.

Исходя из характера арабского текста и с учетом связанных с его формализацией проблем, следует сделать вывод о том, что возможны два критерия, которые могут лежать в основе подсчета и классификации вхождений глагольных словоформ:

1) Выделение и классификация на основе формального критерия графической реализации вхождений глаголов X породы.

2) Выделение и классификация на основе грамматических признаков – т.е. с учетом исходного корня, а также словообразовательной модели и словоизменительных аффиксов.

РЕЗУЛЬТАТЫ

Квантитативный анализ текста Корана продемонстрировал, что в его состав входит 369 глаголов X породы. Определение частотности употребления словоформ глаголов в X породе позволило зафиксировать, что наиболее часто встречаются в тексте формы глагола во мн. ч., м.р., а именно: *yastaṭī'ūna* **يَسْتَطِيعُونَ** «они (м.р.) могут» – 15 вхождений, *yastahzi'ūna* **يَسْتَهْزِئُونَ** «они (м.р.) насмежаются» – 14 вхождений, *istakbarū* **اسْتَكْبَرُوا** «они (м.р.) возгордились» – 11 вхождений, *yastakbirūna* **يَسْتَكْبِرُونَ** «они (м.р.) являются высокомерными» - 7 вхождений и *tasta'ḡilūna* **تَسْتَعْجِلُونَ** «вы (м.р.) торопите» – 6 вхождений соответственно. Например:

yastaṭī'ūna **يَسْتَطِيعُونَ** «они (м.р.) могут»

لَا يَسْتَطِيعُونَ نَصْرَ أَنْفُسِهِمْ وَلَا هُمْ مِمَّا يُصْحَبُونَ

«Разве у них есть боги, которые защитят их от Нас? Не могут они помочь самим себе и не будут от Нас избавлены»¹ (21:44 (43))

yastahzi'ūna يَسْتَهْزِئُونَ «они (м.р.) насмеваются».

وَمَا يَأْتِيهِمْ مِنْ رَسُولٍ إِلَّا كَانُوا بِهِ يَسْتَهْزِئُونَ

«И не приходил к ним никакой посланник, чтобы они над ним не смеялись.» (15:11).

istakbarū - اسْتَكْبَرُوا «они (м.р.) были высокомерны»

فَإِنْ اسْتَكْبَرُوا فَالَّذِينَ عِنْدَ رَبِّكَ يُسَبِّحُونَ لَهُ بِاللَّيْلِ وَالنَّهَارِ

«А если они возгордятся, то те, которые у Господа твоего, прославляют Его и ночью и днем, и они не устают.» (41:38)

قَالَ الَّذِينَ اسْتَكْبَرُوا إِنَّا بِالَّذِي آمَنْتُمْ بِهِ كَافِرُونَ

Сказали те, которые превозносились: "Поистине, мы не веруем в то, во что вы уверовали!" (7:74 (76)).

tasta 'gilūna تَسْتَعْجِلُونَ «они (м.р.) спешили, они ускоряли, торопили».

دُوفُوا فَنَنْتَكُم هَذَا الَّذِي كُنْتُمْ بِهِ تَسْتَعْجِلُونَ

«Вкусите ваше испытание! Это то, с чем вы торопили!» (51:14).

Анализ приведенной статистики демонстрирует тот факт, что наиболее частотные глаголы в десятой порода представлены формами множественного числа мужского рода, что позволяет учесть данный вывод в контексте изучения стилистики текста Корана.

Если рассматривать общее количество вхождений глагола istakbarū – اسْتَكْبَرُوا «они (м.р.) возгордились» с учетом его написаний с союзами – (всего 20

¹ Здесь и далее – цит. по переводу Корана И.Ю. Крачковского [Коран].

вхождений), то данная форма – т.е. перфект 3 л. м. р. мн. ч., производная от корня **كبر** наиболее частотна в тексте Корана среди глаголов X породы.

Таблица 1. Наиболее частотные глагольные словоформы X породы в тексте Корана

№ п.п.	Глагол			Кол-во вхождений без слитных союзов и частиц	Общее кол-во вхождений	Соотношение к общему кол-ву наиболее частотных глаголов (%)
1	yastaṭī'ūna	يَسْتَطِيعُونَ	'они (м.р.) могут'	15	15	20
2	yastahzi'ūna	يَسْتَهْزِئُونَ	'они (м.р.) насмеваются'	14	14	18,66667
3	istakbarū	اسْتَكْبَرُوا	'они (м.р.) возгордились'	11	20	26,66667
4	yastakbirūna	يَسْتَكْبِرُونَ	'они (м.р.) являются высокомер- ными'	7	7	9,333333
5	tasta'gīlūna	تَسْتَعْجِلُونَ	'вы (м.р.) торопите'	6	19	25,33333
Всего	---	---		53	75	100%

Кроме того, количественный анализ глаголов X породы в тексте Корана показал, что наиболее частотными глаголами являются глаголы, производные от корней: **طوع** – 42, **غفر** – 40, **كبر** – 40, **جوب** – 28, **هزأ** – 21, **عجل** – 19, и **أذن** – 12 вхождений соответственно.

Таблица 2. Корни, имеющие наиболее частотные производные глаголов X породы в тексте Корана

№ п.п.	Глагол	Число глаголов X породы, производных от указанного корня	Соотношение к общему количеству глаголов (%)
1	طوع	42	11,38211
2	غفر	40	10,84011
3	كبر	40	10,84011
4	جوب	28	7,588076

№ п.п.	Глагол	Число глаголов X породы, производных от указанного корня	Соотношение к общему количеству глаголов (%)
5	هزأ	21	5,691057
6	عجل	19	5,149051
7	أذن	12	3,252033
Всего	---	202	54,74255

Например:

فَمَا اسْتَطَاعُوا أَنْ يَظْهَرُوهُ وَمَا اسْتَطَاعُوا لَهُ نَقْبًا

«И не могли они взобраться на это, и не могли там продырявить». (18:97)

قَالَ سَلَامٌ عَلَيْكَ سَأَسْتَغْفِرُ لَكَ رَبِّي إِنَّهُ كَانَ بِي حَفِيًّا

Он сказал: «Мир тебе! Я буду просить прощения для тебя у моего Господа: ведь Он ко мне милостив». (19:47)

فَأَسْتَجَابَ لَهُمْ رَبُّهُمْ أَنِّي لَا أُضِيعُ عَمَلَ عَامِلٍ مِنْكُمْ مِنْ ذَكَرٍ أَوْ أُنْثَىٰ

«И ответил им их Господь: «Воистину, Я не пренебрегу ни одним деянием, свершенным кем-либо из вас, мужчиной или женщиной ...»» (3:195)

إِذْ تَسْتَعِينُونَ رَبِّكُمْ فَأَسْتَجَابَ لَكُمْ أَنِّي مُمِدُّكُمْ بِالْفِ مِّنَ الْمَلَائِكَةِ

«И вот, взывали вы за помощью к вашему Господу, и Он ответил вам: «Я поддержу вас тысячью ангелов, следующих друг за другом!»» (8:9)

إِذَا سَمِعْتُمْ آيَاتِ اللَّهِ يُكْفَرُ بِهَا وَيُسْتَهْزَأُ بِهَا فَلَا تَفْعَدُوا مَعَهُمْ

«... что когда вы слушаете знамения Аллаха, в которые не веруют и над которыми издеваются, то не сидите вместе с ними, ...» (4:140)

فَإِنَّ لِلَّذِينَ ظَلَمُوا ذُنُوبًا مِّثْلَ ذُنُوبِ أَصْحَابِهِمْ فَلَا يَسْتَعْجِلُونَ

«Поистине, тем, которые были тиранами, некая доля, подобная доле их сотоварищей, и пусть они не торопят Меня!» (51:59)

لَا يَسْتَأْذِنُكَ الَّذِينَ يُؤْمِنُونَ بِاللَّهِ وَالْيَوْمِ الْآخِرِ أَنْ يُجَاهِدُوا بِأَمْوَالِهِمْ وَأَنْفُسِهِمْ وَاللَّهُ عَلِيمٌ بِالْمُتَّقِينَ

«Не спрашивают у тебя позволения те, которые веруют в Аллаха и в последний день, чтобы сражаться своим имуществом и своими душами. А Аллах знает богобоязненных!» (9:44)

Что касается распределения в зависимости от категорий перфекта или имперфекта, а также числа лица и рода, то из 369 вхождений глаголов X породы в тексте Корана в перфекте отмечено – 124, и в имперфекте – 245 вхождений соответственно.

Таблица 3. Частотность использования грамматических категорий в словоформах X породы в тексте Корана

Перфект				Имперфект		
Лицо/ Число	Ед. ч.	Дв. ч.	Мн. ч.	Ед. ч.	Дв. ч.	Мн. ч.
3 л. м.р.	44 ²	2	55 ³	40 ⁴	1	101 ⁵
3 л. ж.р.	2		0	1	0	1
2 л. м.р.	6		14	36 ⁶	1 ⁷	45 ⁸
2 л. ж.р.	0	0	0	0	2 ⁹	
1 л.	2		5	5		6
Итого	130¹⁰			239^{11 12}		
Всего	369					

Крайне ограниченная сфера применения форм женского рода 2 лица единственного, множественного и двойственного числа может быть косвенным указанием на гендерный состав аудитории проповеди. Впрочем, данный показатель отчасти связан и с тем, что формы двойственного и множественного

² В том числе 4 глагола в страдательном залоге.

³ В том числе 6 глаголов в страдательном залоге.

⁴ В том числе 1 глагол в страдательном залоге

⁵ В том числе 4 глагола в страдательном залоге

⁶ В том числе 24 глагола в повелительном наклонении.

⁷ Глагол в повелительном наклонении.

⁸ В том числе 18 глаголов в повелительном наклонении.

⁹ В том числе 1 глагол в повелительном наклонении.

¹⁰ В том числе 10 глаголов в страдательном залоге.

¹¹ В том числе 44 глагола в повелительном наклонении.

¹² В том числе 5 глаголов в страдательном залоге.

числа мужского рода используются и в случае смешанного состава участников коммуникации.

Залог. Категория страдательного залога представлена в исследуемом тексте десятью вхождениями глаголов в страдательном залоге в перфекте. Из них 6 случаев вхождений глаголов в форме 3 л. мн.ч. м.р. и 4 случая – в форме 3 л. ед.ч. м.р. Представлено также 5 случаев вхождений глаголов в страдательном залоге в имперфекте, в том числе 1 глагол в форме 3 л. ед. ч. м.р. и 4 глагола в 3 л. мн.ч. м.р.

Таблица 4. Вхождения глаголов X породы в страдательном залоге

Глагол		Перевод	Количество вхождений
ustud‘ifū	اسْتَضْعَفُوا	«они были угнетены»	5
ustuhfizū	اسْتَحْفِظُوا	«им было доверено»	1
ustuhzi’a	اسْتَهْزَأَ	«над ним издевались»	3
ustujība	اسْتَجِيبَ	«ему был дан ответ»	1
yustahza’u	يُسْتَهْزَأُ	«над ним издеваются»	1
yustad‘afūna	يُسْتَضْعَفُونَ	«их считали слабыми»	1
yusta‘tabūna	يُسْتَعْتَبُونَ	«не будет им оказано милости» ¹³	3

Время. Рассмотрение количественного состава показывает, что соотношение глаголов в перфекте и имперфекте составляет 136 к 239. Другими словами, в 64,77 % случаев нарратив главным образом относится к настоящему или будущему времени, и лишь в 35,23 % случаев имеет соотнесенность с событиями в прошлом.

Императив. В тексте Корана представлено 44 случая вхождений форм императива X породы глагола, или около 11,92 % от общего числа глаголов X породы.

Таблица 5. Глаголы X породы в императиве в тексте Корана

Число	Единственное число		Двойственное число		Множественное число		Всего
	м.р.	ж.р.	м.р.	ж.р.	м.р.	ж.р.	
Род	м.р.	ж.р.	м.р.	ж.р.	м.р.	ж.р.	
Количество вхождений	24	1	1		18	-	44

¹³ См. Коран перевод И.Ю.Крачковского.

Глаголы X породы среди глаголов других пород в тексте Корана. В целом, количество глаголов в X породе составляет 1,9 % от общего числа – 19356 – глаголов, встречающихся в тексте Корана.

Таблица 6. Процентное соотношение вхождений глаголов в каждой из пород

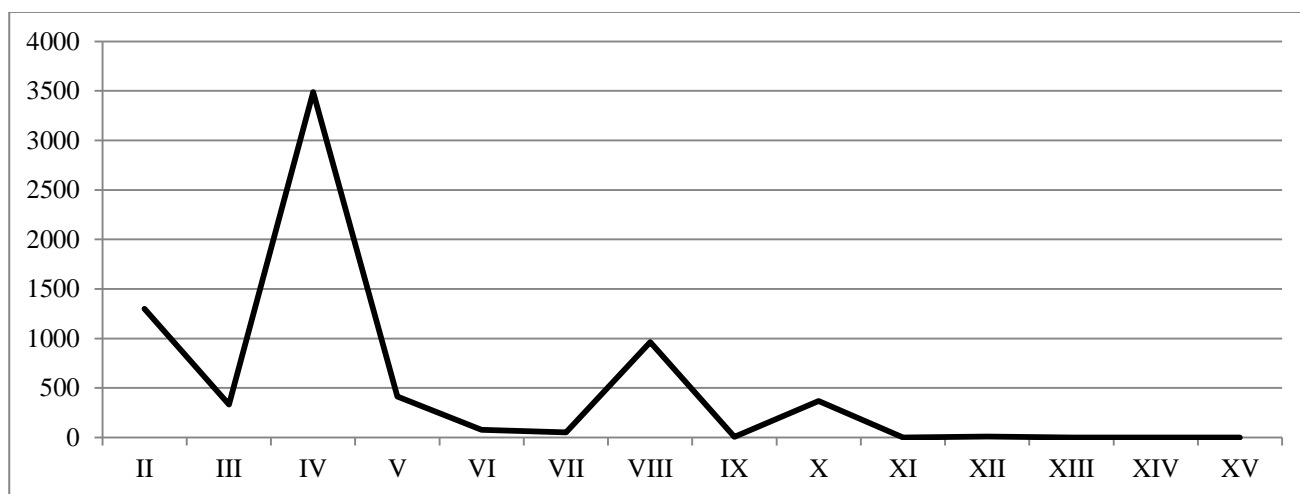
№ п.п.	Породы глаголов в Коране	Число глаголов в каждой из пород	Процентное соотношение для вхождений глаголов каждой из пород в тексте Корана (в %)
1	I	12347	63,78901
2	II	1300	6,716264
3	III	334	1,725563
4	IV	3487	18,01509
5	V	414	2,138872
6	VI	77	0,397809
7	VII	51	0,263484
8	VIII	963	4,975201
9	IX	5	0,025832
10	X	369	1,906386
11	XII	9	0,046497
Итого	11	19356	100 %

Полученные данные показывают, что общее количество глаголов в десятой породе (369 вхождений) сопоставимо с их количеством в пятой и третьей породах (414 и 334 вхождения соответственно). Данный показатель значительно выше шестой породы – 77, седьмой – 51, девятой – 5, и двенадцатой породе (9 вхождений), хотя и заметно меньше, чем в первой – 12347, второй – 1300, четвертой – 3487 и восьмой (963 вхождений) породах. Данное обстоятельство во многом, вероятно, связано с семантикой, передаваемой глаголами X породы в тексте Корана.

На приведенное ниже диаграмме демонстрируется частотность реализации глаголов в производных породах в тексте Корана. Учитывая, что общее количество вхождений глаголов первой породы демонстрирует максимально высокий показатель, он не включен в данный график.

В дальнейшем планируется сопоставление приведенных данных с фактами арабского литературного языка, полученными на основе проведения корпусных исследований.

График 1. Частотность употребления глаголов в производных породах в тексте Корана



ВЫВОДЫ

Автоматические методы обработки данных позволяют проводить текстологические исследования на качественно новом уровне, что особенно актуально при изучении особенностей языка Корана.

Несмотря на наличие значительного количества исследований, связанных с разработкой и применением информационных технологий для контент-анализа арабографических документов, имеющиеся решения не всегда подходят для лингвистических задач теоретического характера. Данное обстоятельство в значительной мере обусловлено тем, что существующие лингвистические программные приложения (машинный перевод и другие) не всегда могут быть использованы для решения локальных проблем. В этой связи цифровая компетентность исследователя-лингвиста позволяет эффективно применять стандартное программное обеспечение для решения задач частного характера.

Разработка алгоритмов процессинга коранического текста с целью определения особенностей функционирования глаголов в X породе позволяет сделать выводы относительно грамматики текста Корана, включая частотность функционирования различных корней и их производных с учетом основных грамматических категорий глагола. Статистические результаты исследования формируют наглядную картину реализации соответствующих моделей, что позволит в дальнейшем провести сопоставительный анализ с фактами арабского литературного языка.

Результаты исследования, с одной стороны, открывают новые теоретические возможности изучения арабского языка в диахронии, а с другой – позволяют применить апробированные решения автоматической обработки

арабографического контента при анализе особенностей других морфологических моделей в тексте Корана, как именных, так и глагольных. Данный материал позволяет также создать и эффективно апробировать алгоритмы автоматизированного анализа текста с возможностью их дальнейшего совершенствования и использования при анализе больших объемов лингвистического материала.

ЛИТЕРАТУРА

1. Svensson, J. (2009). Computing Qur'ans: A Suggestion for a Digital Humanities Approach to the Question of Interrelations between English Qur'an Translations. *Islam and Christian-Muslim Relations*, 30 (2), 211–229.
2. Резван Е.А. Коран и его мир. СПб.: Петербургское востоковедение, 2001. 608 с.
3. The Quranic Arabic Corpus. (n.d.). Retrieved from <http://corpus.quran.com/>
4. Atwell, E. (2015). Quran Analysis. Retrieved from <https://qurananalysis.com/info/index.php?page=about>
5. (n.d.). *Dictionary of Qur'anic Usage Online*. Leiden, The Netherlands: Brill. Retrieved from <https://brill.com/view/db/dquo>
6. Badawi, E. M., & Abdel Haleem, M. (Eds). (2008). *Arabic - English Dictionary of Qur'anic Usage*. Leiden, The Netherlands: Brill.
7. Кудрявцева А. Ю., Резван Е. А. Лексико-грамматический строй Корана: Учебно-методический корпус. [Электронный ресурс]. URL: http://www.kunstkamera.ru/museums_structure/research_departments/islam/corpuscoranicum_petropolitana/koran (дата обращения: 10.07.2020).
8. Wright, W. A. (1896). *Grammar of the Arabic Language*. (3rd ed., vol. 1). Cambridge: at the University Press.
9. Берникова О. А., Редькин О. И. Язык Корана: особенности морфологии глагола // Вестник РФФИ. Гуманитарные и общественные науки. 2019. № 4 (97). С. 59-72.
10. Коран. Перевод и комментарии И.Ю.Крачковского. Под редакцией В.И.Беляева. Изд. 2-е. М.: Главная редакция восточной литературы издательства «Наука», 1986. 728 с.

BIBLIOGRAPHIC REFERENCES

1. Svensson, J. (2009). Computing Qur'ans: A Suggestion for a Digital Humanities Approach to the Question of Interrelations between English Qur'an Translations. *Islam and Christian-Muslim Relations*, 30 (2), 211–229.
2. Rezvan, E. A. (2001). *Koran i ego mir* [The Quran and its World]. Saint-Petersburg, Russia: Peterburgskoye Vostokovedenie. 608 p. (In Russian)
3. The Quranic Arabic Corpus. (n.d.). Retrieved from <http://corpus.quran.com/>
4. Atwell, E. (2015). Quran Analysis. Retrieved from <https://qurananalysis.com/info/index.php?page=about>
5. (n.d.). Dictionary of Qur'anic Usage Online. Leiden, The Netherlands: Brill. Retrieved from <https://brill.com/view/db/dquo>
6. Badawi, E. M., & Abdel Haleem, M. (Eds). (2008). *Arabic - English Dictionary of Qur'anic Usage*. Leiden, The Netherlands: Brill.
7. Kudryavtseva, A. Yu., & Rezvan, E. A. Leksiko-grammaticheskiy stroy Korana. Uchebno-metodicheskiy korpus [Lexico-grammatical structure of the Qur'an: Educational-methodical corpus]. Retrieved from http://www.kunstkamera.ru/museums_structure/research_departments/islam/corpuscoranicum_petropolitana/koran (In Russian)
8. Wright, W. A. (1896). *Grammar of the Arabic Language*. (3rd ed., vol. 1). Cambridge: at the University Press.
9. Bernikova, O. A., & Red'kin, O. I. (2019). Yazyk Korana: osobennosti morfologii glagola [The language of the Quran: features of the morphology of the verb]. *Vestnik RFFI. Gumanitarnyye i obshchestvennyye nauk* [RFBR Bulletin. Humanities and Social Sciences]. 4 (97), p. 59-72. (In Russian)
10. *Koran. Perevod i kommentarii I. Yu. Krachkovskogo* [Quran. Translation and commentaries by I. Yu. Krachkovskiy]. (1986). V.I. Belyaev (Ed.) (2nd ed.). Moscow, Russia: Glavnaya redaktsiya vostochnoy literatury izdatel'stva «Nauka» [The main editorial office of the oriental literature of Nauka publishing house]. (In Russian)

Благодарность: Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-012-00849.

Acknowledgments: *The reported study was funded by RFBR according to the research project № 18-012-00849.*

Информация об авторах

Профессор, д.фил.н.,
Олег Иванович Редькин
Кафедра арабской филологии
Научная лаборатория анализа и
моделирования социальных процессов
Санкт-Петербургский
государственный
Университет
199034, Санкт-Петербург
Университетская наб. 11
Российская Федерация
o.redkin@spbu.ru
WOS Research ID: J-8892-2013
Scopus Author ID: 55355596600

Доцент, к.фил.н.,
Ольга Александровна Берникова
Кафедра арабской филологии
Научная лаборатория анализа и
моделирования социальных процессов
Санкт-Петербургский
государственный
Университет
199034, Санкт-Петербург
Университетская наб. 11
Российская Федерация
o.bernikova@spbu.ru
WOS Research ID: H-2273-2013
Scopus Author ID: 55614667800

Information about authors

Professor, Doctor of Science,
Oleg Ivanovich Redkin
Arabic Philology Department
Research Laboratory for Analysis and
Modeling of Social Processes
Saint-Petersburg State University
199034, Saint-Petersburg
11 Universitetskaya emb.
Russian Federation
o.redkin@spbu.ru
WOS Research ID: J-8892-2013
Scopus Author ID: 55355596600

Associate Professor, PhD
Olga Aleksandrovna Bernikova
Arabic Philology Department
Research Laboratory for Analysis and
Modeling of Social Processes
Saint-Petersburg State University
199034, Saint-Petersburg
11 Universitetskaya emb.
Russian Federation
o.bernikova@spbu.ru
WOS Research ID: H-2273-2013
Scopus Author ID: 55614667800

Раскрытие информации о конфликте интересов: *Автор заявляет об отсутствии конфликта интересов.*

Conflicts of Interest Disclosure: *The author declares Conflicts of Interest Disclosure.*