

PAŁAEOBULGARICA

СТАРО-  
БЪЛГАРИСТИКА

BULGARIAN ACADEMY OF SCIENCES  
CYRILLO-METHODIAN RESEARCH CENTRE

БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ  
КИРИЛО-МЕТОДИЕВСКИ НАУЧЕН ЦЕНТЪР

4

THIRTY-EIGHTH YEAR • 2014 • ГОДИНА ТРИДЕСЕТ И ОСМА

Д-р. Славия БЪРЛИЕВА (България), prof. Charles GRIBBLE (USA), проф. Клементина ИВАНОВА (България), prof. Roland MARTI (Deutschland), проф. Ангелина МИНЧЕВА (България), д-р. Татяна МОСТРОВА (България), д-р. Георги НИКОЛОВ (България), проф. Светлана НИКОЛОВА (България) – главен редактор, проф. Георги ПОПОВ (България), Галина ФИЛИПОВА – секретар

*Издаването на настоящия брой от списанието е с финансовата подкрепа на Фонд „Научни изследвания“ при Министерството на образованието и науката.*

Предпечатна подготовка А. Борисова  
Коректор М. Евлимова

Печатни ходи 7,25 Формат 700 x 1000/16

Печатница „Симолни-94 СД“  
София, бул. „Владимир Вазов“ 15

Адрес на редакцията: 1000 София, ул. „Московска“ 13  
за кореспонденция: 1000 София, п.к. 432  
e-mail: kmpc@bas.bg; kmpc@abv.bg

# PAEAEOBVJΓARICA

## СТАРО- БЪЛГАРИСТИКА

БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ / BULGARIAN ACADEMY OF SCIENCES  
КИРИЛО-МЕТОДИЕВСКИ НАУЧЕН ЦЕНТЪР/CYRILLO-METHODIAN RESEARCH CENTRE

XXXVIII (2014), 4

### СЪДЪРЖАНИЕ / СОДЕРЖАНИЕ / CONTENTS

Нели Ганчева, Галина Филипова (София, България) – Славянското средновековно писмено наследство и информационните технологии .....	3
СТАТИИ / СТАТЬИ / ARTICLES	
Тимофей Архангелский, Екатерина Мишина, Анна Пичхадзе (Москва, Россия) – Система электронной грамматической разметки древнерусских и церковнославянских текстов .....	21
Emanuel Wenger (Vienna, Austria), Elena Ukhanova (Moscow, Russia) – An Internet Portal for Watermarks .....	38
Наталья Нечунаева (Таллин, Эстония), Алексей Нечунаев (Берлин, Германия) – Рукописи славянских миней XI–XIV вв. и принципы их классификации .....	45
Мария Малыгина (Москва, Россия) – К вопросу о ритмической организации славянских гимнографических текстов .....	57
Магдалена Абаджиева (София, България) – Книжовен език на народна основа ли е павликанската книжнина от втората половина на XVIII век? .....	68

### ПРЕГЛЕД / ОБЗОРИ И РЕЦЕНЗИИ / SURVEY AND BOOK REVIEWS

Мая Иванова, Евелина Джевиевца (София, България) – Юбилейна конференция „100 години кирилometодиевистика в България“ .....	79
Златка Тыменова, Георги Христовски (Лисабон, Португалия) – От мисия към мисианство? .....	94
Маргарита Младенова (София, България) – Увлекателен преглед на кирило-методиевската проблематика на фона на взаимоотношенията между Византия и Рим .....	103
Иван Илиев (Кърджали, България) – Принос към проучването на среднобългарския превод на словата на книжовника авва Дорогей от Газа .....	111

*Наталья Нечунаева (Таллинн, Эстония), Алексей Нечунаев (Берлин, Германия)*

## РУКОПИСИ СЛАВЯНСКИХ МИНЕЙ XI–XIV ВВ. И ПРИНЦИПЫ ИХ КЛАССИФИКАЦИИ

### 1. ВВЕДЕНИЕ

Минея – богослужебная книга, содержащая тексты, прославляющие святых и церковные праздники, которые расположены по дням каждого месяца. В Минее, памятнике многосписочном, в процессе бытования возникло большое количество разночтений на текстовом и языковом уровнях. Отсюда – важность классификации памятника, представленного, согласно Сводному каталогу, 68 рукописями XI–XIII вв.<sup>1</sup> Самое большое количество списков этого периода имеет Минея на май – одиннадцать.

Традиция классификации гимнографических текстов восходит к работам И. В. Ягича<sup>2</sup>. Лингвотекстологические параметры для классификации списков и выделения типов памятника: состав памятника (набор памятей), структура службы (порядок песнопений) и особые лингвистические чтения. Из современных исследователей Минеи, придерживающихся такого же подхода, необходимо назвать Марию Йовчеву, описывающую служебную Минею на апрель Соф. 199 XII–XIII вв. из собрания РНБ<sup>3</sup> с учетом ка-

<sup>1</sup> Сводный каталог славяно-русских рукописных книг, хранящихся в СССР XI–XIII вв. Москва, 1984 (далее – СК).

<sup>2</sup> Ягич, И. В. Служебные минеи за сентябрь, октябрь и ноябрь в церковно-славянском переводе по русским рукописям 1095–1097 годов. СПб., 1886, с. LI–LXXII, LXXVII–LXCVI.

<sup>3</sup> Здесь и далее в статье использованы сокращения:

Рукописные хранилища: РНБ – Российская национальная библиотека, ГИМ – Государственный исторический музей, БАН – Библиотека Академии наук, РГАДА – Российский государственный архив древних актов.

Источники: РНБ Соф. 202 – Минея служебная, май (Путятина Минея – ПМ), XI в.; Соф. 203 – Минея служебная, май, XII в.; Соф. 204 – Минея служебная, май, XIII в.; Q.п.I.25 – Минея праздничная, май, кон. XII в. – нач. XIII в.; Соф. 382 – Минея праздничная, май, кон. XIII в. – нач. XIV в.; Соф. 199 – Минея служебная, апрель, XII в.; РНБ греч. 552 – Минея греческая на март–август, XII в.; ГИМ Син. 166 – Минея служебная, май, XII в.; БАН – БАН 4.5.10 Минея служебная, май, XIII в.; БАН 16.14.13 Минея служебная, май, нач. XIV в.; РГАДА ф. 381 (Син. Тип.) № 113 – Минея служебная, май, XIII в. (Т. 113); РГАДА ф. 381 (Син. Тип.) № 112 – Минея служебная, май, XIV в. (Т. 112).

лендарных особенностей рукописи (набор памятей), состава службы (порядок песнопений), морфологических, лексических и синтаксических различий<sup>4</sup>.

Текстологическая типология и лингвистические чтения позволяют объединить списки Минеи на май в 4 типа:

1. Путятинская Минея и ее аналоги, принадлежащие к „архаичному“ типу Минеи. Возможно, что такой порядок песнопений как в ПМ был еще до введения Студийско-Алексеевского устава;

2. Списки студийского типа XII–XIII вв. – основной корпус СК;

3. Списки раннего иерусалимского типа, описанные в Предварительном списке<sup>5</sup>, например, рукопись Минеи на май Т.113 XIV в.;

4. Списки иерусалимского типа, в большинстве случаев, зафиксированные в каталогах хранилищ, датируемые XIV–XVII вв. Например, рукопись KB Rålamb 4: 0 n:0 130 Минея праздничная XVII в. из собрания Королевской библиотеки Швеции (Стокгольм).

Состав памятника и структура службы – устойчивые показатели для типологической характеристики списка: набор памятей и порядок песнопений в службе связаны с определенным уставом. По набору памятей списки XI–XIV вв. второго типа из СК соответствуют, в основном, Студийскому уставу, списки третьего и четвертого типа – Иерусалимскому уставу. В списках Минеи отмечены три возможных варианта порядка следования песнопений: 1) канон – стихиры – седален; 2) седален – стихиры – канон; 3) стихиры – канон, прерываемый после третьей песни седалном и после шестой песни кондаком и икосом. Первые два варианта расположения песнопений жанровые, третий соответствует порядку следования при службе. Первый вариант является характерной чертой списков типа, названного нами „архаичным“. Коллекция таких списков в настоящее время состоит из 4-х древнерусских списков, включая Путятинскую Минею, 5-ти южнославянских и 1-го древнегреческого списка<sup>6</sup>. Второй вариант соотносится со Студийским уставом, третий – с Иерусалимским.

Классификация может быть более подробной – по некоторым параметрам среди студийских списков выделяется майская Минея Соф. 203 XII в. Текст рукописи показывает, что в пределах большого массива однотипных списков студийского типа возможно существование нестандартного подтипа. Список XII в. Соф. 203 имеет текстологические особенности, связанные

<sup>4</sup> Й о в ч е в а, М. Старобългарският служебен миней. София, 2014, 231–277.

<sup>5</sup> Предварительный список славяно-русских рукописей XI–XIV вв., хранящихся в СССР. Археографический ежегодник за 1965 г. Москва, 1966 (далее – ПС).

<sup>6</sup> Н е ч у н а е в а, Н. Минея: История текста и списков XI–XIV вв. – В: Славянское языкознание: покидая XX век. К XIV Международному съезду славистов (Охрид, 10–16.09.2008). Tartu, 2008, 9–23.

с набором памятей по иерусалимскому типу и особыми чтениями, например, реставрацией грецизмов в тексте: олъвемь вместо масломь<sup>7</sup>.

Названные параметры в разной мере учитываются исследователями гимнографии, которые следуют текстологическим традициям, сложившимся при описании книг Ветхого и Нового завета<sup>8</sup>. Однако эти методы классификации не базируются на целостном использовании информации, доступной в списках Минеи. Поэтому в данной статье предлагается использовать методы классификации, основанные на свойствах данных, в данном случае, на информации о разночтениях разного плана, содержащейся в списках Минеи. Источник материала – списки первого и второго типа, в которых фиксируются языковые разночтения – лексические и грамматические. Фонетические различия на данном этапе анализа не учитываются.

В статье используется новый подход к классификации и анализу данных текстов, который мы заимствуем из литературы, посвящённой машинному обучению и информационному поиску. Сходные и различные списки из общей группы миней первого и второго типа предполагается найти с помощью математической модели. Для формального анализа списков миней строится векторная модель, где эти списки представлены векторами в общем пространстве. Посредством векторной модели можем найти похожие и отличные списки и упорядочить списки миней первого и второго типа в группы. При анализе учитываем лексические, а также лексические и грамматические различия. Группы, сформированные в результате статистического анализа, являются наборами наиболее схожих списков миней внутри определенных текстологических типов на уровне лексики, а также лексики и грамматики.

Данная статья показывает, что методы информационного поиска могут быть успешно применены к анализу древних текстов. Полученные результаты позволяют выявить схожие и различные списки Минеи и соотнести результаты с имеющимися в литературе. Дальнейшая часть статьи организована следующим образом. В части 2 описываются миней и фрагменты текстов, использованные в данной статье. В части 3 строится векторная модель. Детальный анализ результатов, полученных с помощью векторной модели, представлен в части 4. Последняя часть является заключительной и обобщает полученные результаты.

<sup>7</sup> Нечунова, Н. Рукопись Минеи Соф. 203 XII в. из собрания РНБ как текст и как список. – В: Лингвокультурное пространство современной Европы через призму больших и малых языков. К 70-летию профессора А. Д. Дуличенко (= Slavica Tartuensia, 9). Тарту, 2011, 167–176.

<sup>8</sup> Алексеев, А. А. Текстология славянской Библии. СПб., 1999, 60–62.

## 2. МИНЕИ

Для анализа с использованием векторной модели привлечено девять списков Миней на май, XI–XIV вв.: семь учтены в СК и два – в ПС. Краткая характеристика списков и места хранения приведены в Таблице 1.

Рукопись	№ в СК	№ в ПС	Место хранения	Шифр в статье
Миней служебная, май, XI в. Пугятин Миней	21	10	РНБ, Соф. 202	ПМ
Миней праздничная, май, отрывок. Кон. XII – нач. XIII в.	156	72	РНБ, Q.п.I.25	Q.п.I.25
Миней служебная, май, XII в.	90	107	РНБ, Соф. 203	Соф.203
Миней служебная, нотированная, май, XII в.	89	106	ГИМ, Син. 166	Син.166
Миней служебная, май, XIII в.	282	436	РНБ, Соф. 204	Соф.204
Миней служебная с проложными чтениями, май–июнь, XIII в.	281	313	БАН, 4.5.10	БАН4.5.10
Миней праздничная. Кон. XIII – нач. XIV в.	454	434	РНБ, Соф. 382	Соф.382
Миней служебная, май, В ПС XIII/XIVв. В БАН нач. XIV в.		435	БАН, 16.14.13	БАН16.14.13
Миней служебная, май, XIV в.		1062	РГАДА, ф. 381 Син. Тип. №112	T.112

Таблица 1. Списки Миней XI–XIV вв.

Чтобы представить списки памятника в векторной форме, были использованы два первых тропаря из стихирь  $\overline{\text{стн}} \cdot \overline{\text{гла}} \cdot \overline{\text{д}} \cdot \overline{\text{по}} \cdot \overline{\text{дасть}}$  знаменне, которая в ПМ и Q.п.I.25 следует после канона службы 21 мая св. царям Константину и Елене, а в остальных списках предшествует канону. Количество слов в двух тропарях каждого списка варьируется от 67 до 59. Покажем в качестве образца первый тропарь по трем рукописям: списки ПМ и Q.п.I.25 по текстологическим параметрам относятся к архаичному типу (других рукописей этого типа среди майских миней нет) и список БАН16.14.13 – к студийскому (один из списков данного типа):

ПМ:

Ороужне крѣпко • црю тѣ дасть нашему • крѣ|стѣ свои чѣстнѣи • нмъ же црѣ|ткова  
на землн| правдѣно • снѣвѣ вѣрою • н црю нѣсѣ съподобн|са • млсрднѣемьсн • съ ннмн  
же тн млств' но съ|мотренне прослав' лаемъ • ѿсе млстнѣи • спсе| дшѣ нашнхъ ÷ (л. 88)

Q.п.I.25:

Оржжне крѣпко цревн дасть • | нашемъ крѣсть твои чѣстнѣи: нмъ же црѣ|ствова: на  
земн праведн: снѣвѣ вѣроѣ: црѣ|ства нѣсѣ (...) съподобн: млсрѣднѣемьсн: съ| ннмъ же  
млсрдѣное: ѿмотренне тѣе слаѣмъ: ѿсе въ|семѣе:: (л. 1)



БАНІ6.14.13:

Ороуѣжыє крѣпко • | црѣвн нашемоу дасть крѣтъ твоѣ чѣнзѣн • н|м же црѣтвова на землѣн •  
правдѣно снаѣа вѣрою | црѣтва нѣнаго сподобн сл • мѣрднѣмъ |сн • с ннмъ же тн  
члѣколюбьноу • съ|мотреннѣ славнмъ • нсе мѣтнвзѣн спѣз | дшмъ нашнм ∴ (л. 50)

### 3. ВЕКТОРНАЯ МОДЕЛЬ

Для статистического анализа мы используем векторную модель и представляем тексты векторами в общем для всей коллекции списков векторном пространстве<sup>9</sup>. Список в векторном пространстве – это вектор весов терминов. В нашем случае термин может быть как единичным словом так и словосочетанием. Более формально:

$D = \{d_1, \dots, d_n\}$  – множество (корпус) документов

$T = \{t_1, \dots, t_m\}$  – множество слов, содержащихся в документах, или словарь.

Следующим шагом является определение веса термина в документе. Вес – это важность термина для идентификации данного списка. Векторным представлением документа и является вектор, который содержит веса каждого термина в документе. Если термин отсутствует в документе, его вес равен нулю. А именно документ  $i$ :

$W(i) = \{w_{1i}, w_{2i}, \dots, w_{mi}\}^T$ , где  $w_{ji}$  – вес термина  $j$  в документе  $i$ .

Существует несколько стандартных способов определения функции взвешивания. В данной статье используется *tf-idf* (term frequency – inverse document frequency, частота термина – обратная частота документа) функция взвешивания<sup>10</sup>. Функция взвешивания рассчитывается следующим образом:

*tf* (term frequency) оценивает важность термина  $t_i$  в конкретном документе  $d_j$

$tf(t_i, d_j) = n_{ij}$ ,  $n_{ij}$  – число вхождений термина в документ  $d_j$ .

*idf* (inverse document frequency) – это инверсия частоты, с которой некоторое слово встречается в документах корпуса. Рассчитывается этот показатель следующим образом:

$$idf(t_i, D) = \log_{10} \frac{|D|}{|d \in D : t_i \in d|}$$

<sup>9</sup> Salton, G., A. Wong, C. S. Yang. A vector space model for automatic indexing. – Communications of the ACM, 18, Nov. 1975, № 11, 613–620; D ub in, D. The Most Influential Paper Gerard Salton Never Wrote. – Library Trends, 52, 2004, № 4, 748–764.

<sup>10</sup> Rajaraman, A., J. D. Ullman. “Data Mining”. Mining of Massive Datasets. Cambridge, 2011, 1–17, doi:10.1017/CBO9781139058452.002.

где  $|D|$  – количество документов, а  $|d \in D : t_i \in d|$  – количество документов, где присутствует термин  $t_i$ .

Вес tf-idf вычисляется как:

$$w_{ij} \equiv tfidf(t_i, d_j, D) = tf(t_i, d_j)idf(t_i, D).$$

Таким образом каждый документ теперь представлен в виде вектора весов в пространстве  $\mathbb{R}^m$ . Это позволяет вычислить схожесть и различие векторов, используя определённые метрики. В данной статье мы рассчитываем Евклидову дистанцию между списками и косинус-сходство списков. Первый объект позволяет выявить наиболее различные списки, второй – наиболее схожие.

Дистанция рассчитывается по формуле:

$$dist(d_i, d_j) = \sqrt{\sum_{k=1}^m (w_{ki} - w_{kj})^2}$$

Чем меньше дистанция, тем ближе списки.

Косинус-сходство списков определяется как:

$S(d_i, d_j) = \tilde{W}(i)^T \tilde{W}(j)$ , где  $\tilde{W}(i)$  – нормализованный по длине вектор  $W(i)$  (использована  $L^2$  норма). Чем больше значение  $S(d_i, d_j)$ , тем больше похожи списки  $d_i$  и  $d_j$ .

#### 4. АНАЛИЗ ТЕКСТОВ

Для анализа с использованием векторной модели привлечено два фрагмента из стихир по девяти спискам Минеи, описанным в Таблице 1. Проанализирована по отдельности схожесть списков по лексике и по лексике и грамматике. Для этого были созданы два словаря, в первом случае словарь содержит 70 терминов, во втором – 79 терминов. В первом случае учитываются лексические различия типа: *сложъ – оугодьнкоу* и *мств'но – мств'ное – члвколювьнюк*. Во втором случае лексические и грамматические различия типа *снавъ – възснавъ – снвѧ*.

Разделение анализа на лексический и лексический и грамматический оказалось важным, поскольку позволило проследить собственно лексическую вариативность и ее связь с определенными списками. Лексические и грамматические изменения в пределах одной языковой единицы показали динамику некоторых грамматических изменений в диахронии и их связь с лексическими изменениями.



Рассмотрим в начале дистанцию между списками, в случае, когда учитывались лексические различия. Данные по этой дистанции представлены в Таблице 2. Общая дистанция показывает, что ПМ и Q.п.I.25 имеют наибольшее удаление от остальных списков. Наименьшая общая дистанция наблюдается у БАН16.14.13. Это объясняется тем, что дистанция между списками Соф.203, Син.166, Соф.204, БАН4.5.10, Соф.382, Т.112 и БАН16.14.13 наименьшая.

	ПМ	Q.п.I.25	Соф.203	Син.166	Соф.204	БАН4.5.10	Соф.382	БАН16.14.13	Т.112
ПМ	0	0,094	0,079	0,079	0,079	0,084	0,079	0,075	0,081
Q.п.I.25	0,094	0	0,078	0,078	0,081	0,084	0,078	0,075	0,072
Соф.203	0,079	0,078	0	0,035	0,043	0,046	0,034	0,026	0,039
Син.166	0,079	0,078	0,035	0	0,044	0,046	0,035	0,025	0,039
Соф.204	0,079	0,082	0,043	0,044	0	0,046	0,043	0,036	0,047
БАН 4.5.10	0,084	0,084	0,046	0,046	0,046	0	0,046	0,039	0,049
Соф.382	0,079	0,078	0,034	0,035	0,043	0,046	0	0,025	0,039
БАН16.14.13	0,075	0,075	0,026	0,025	0,036	0,039	0,025	0	0,030
Т.112	0,081	0,072	0,039	0,039	0,047	0,049	0,039	0,030	0
Общая дистанция	0,653	0,645	0,84	0,383	0,422	0,442	0,382	0,332	0,399

Таблица 2. Евклидова дистанция между списками (лексика)

Наименьшее значение для каждого списка в столбце выделено курсивом. Наименьшее значение в последней строке, показывающую общую дистанцию между списками, выделено курсивом.

Данные косинус-сходства списков в случае, когда учитывались лексические различия, представлены в Таблице 3. В принципе, изменение меры длины даёт схожие результаты.

	ПМ	Q.п.I.25	Соф.203	Син.166	Соф.204	БАН4.5.10	Соф.382	БАН16.14.13	Т.112
ПМ	1	0,227	0,076	0,091	0,159	0,069	0,079	0,157	0,079
Q.п.I.25	0,227	1	0,078	0,075	0,060	0,057	0,082	0,128	0,305
Соф.203	0,076	0,078	1	0,262	0,183	0,171	0,289	0,486	0,242
Син.166	0,091	0,075	0,262	1	0,170	0,162	0,269	0,505	0,235
Соф.204	0,159	0,060	0,183	0,170	1	0,331	0,187	0,312	0,152
БАН 4.5.10	0,069	0,057	0,171	0,162	0,331	1	0,173	0,342	0,175
Соф.382	0,079	0,082	0,289	0,269	0,187	0,173	1	0,482	0,247
БАН16.14.13	0,157	0,128	0,486	0,505	0,312	0,342	0,482	1	0,485
Т.112	0,079	0,305	0,242	0,235	0,152	0,175	0,247	0,485	1
Общая схожесть	1,941	2,018	2,789	2,841	2,550	2,514	2,801	3,938	2,955

Таблица 3. Косинус-сходство списков (лексика)

Наибольшее значение для каждого списка в столбце выделено курсивом. Наибольшее значение в последней строке, показывающую общую схожесть списков, выделено курсивом.

Проанализируем лексику, опираясь на Таблицы 2 и 3.

Списки ПМ и Q.п.I.25 отстоят от других списков по лексическим чтениям: мѣств'но – мѣсрѣд'ное (Q.п.I.25) – члвколювьное (студийские списки). Подобные различия могут быть связаны с греческими чтениями. Лексема члвколювьное – греческая калька φιλόανθρωπος, допускающая в славянском тексте и другой перевод – мнлостнвнзгн: зѣло сѣддѣко н мнлостнвнзгн (φιλόανθρωπος) Гр. Наз. XI в.<sup>11</sup> Это различие представлено в ПМ в виде мѣств'но. Второй член в различии мѣсрѣд'ное может быть как задачей греческого φιλόανθρωπος, так и греческого εὐσπλαγχνος 'мнлосрѣдз'<sup>12</sup>. Лексические различия блговѣрннчмоу – блгочестнвнчмоу восходят к одной греческой лексеме εὐσεβής<sup>13</sup>. Их варьирование возникло на славянской почве. Такого же происхождения различия слоузѣ (ПМ, Q.п.I.25) – оугодннкоу (списки студийского типа).

Полученные результаты говорят о том, что ПМ наиболее схожа с Q.п.I.25, а Q.п.I.25 в свою очередь с Т.112: твѣ/тво (только в этих двух списках). В греческом тексте местоимение отсутствует. Появление связано с двумя списками, возможно, перевод был сделан с другого греческого списка.

Наибольшее общее сходство наблюдается у БАН16.14.13 со всеми списками: свон (ПМ) – твон (остальные списки). Все списки, кроме ПМ, дают соответствие оригиналу: τὸν σταυρόν σου τὸν τίμιον. Отсутствие лексемы в ПМ свидетельствует о наличии другого оригинального текста. Различие прослав'ляемз (ПМ) – славнмз подтверждает это наблюдение. Различие восходит к греческому ἐνδοξάζειν – δοξάζειν 'прославнтн–славнтн'. В греческом тексте (греч. 552) представлена лексема δοξάζομεν, которая, как и в первом случае, передана во всех списках, кроме ПМ.

Список БАН16.14.13 очень схож со списками Соф.203, Син.166, Соф.204, БАН4.5.10, Соф.382 и Т.112. Наиболее близкими являются БАН16.14.13 и Син.166. Лексема крѣпкко функционирует во всех списках, кроме Соф.382, в котором неподнмо – искажение лексемы неповѣднмо и Т.112, в котором представлено грамматическое различие исходной лексемы в форме крѣпккоу. Совпадение в списках второго типа прилагательного происходит в краткой форме крѣпкко.

Таким образом можно выделить списки ПМ и Q.п.I.25, которые отстоят от других по лексическому сходству. Косинус-сходство указывает, что эти рукописи схожи между собой. Список БАН16.14.13 напротив имеет наибольшее сходство с остальной группой исследуемых списков. Предыдущие исследования, использующие другую методику, приходили к выводу о том, что ПМ и Q.п.I.25 на основании текстологического и лингвистического сходства выделяются в особый тип – архаический. Наибольшее сходство списка БАН16.14.13 с остальной группой показал только статистический анализ.

<sup>11</sup> Срезневский, И. И. Словарь древнерусского языка. Т. 2. Москва, 1989, с. 139.

<sup>12</sup> Старославянский словарь (по рукописям X–XI веков). Москва, 1999, с. 326.

<sup>13</sup> Срезневский, И. И. Словарь древнерусского языка. Т. 1. Москва, 1989, с. 92, 109, 403.

Данные по дистанции между списками для случая, когда учитывались лексические и грамматические различия, представлены в Таблице 4. Общая дистанция показывает, что ПМ и Q.п.I.25 имеют наибольшее удаление от остальных списков. Наименьшая общая дистанция наблюдается у БАН16.14.13. Эти результаты схожи с полученными ранее результатами, когда учитывались только лексические разночтения.

	ПМ	Q.п.I.25	Соф.203	Син.166	Соф.204	БАН4.5.10	Соф.382	БАН16.14.13	T.112
ПМ	0	0,093	0,077	0,076	0,083	0,082	0,077	0,076	0,074
Q.п.I.25	0,093	0	0,070	0,070	0,080	0,076	0,071	0,070	0,067
Соф.203	0,077	0,070	0	0,030	0,049	0,042	0,032	0,030	0,037
Син.166	0,076	0,070	0,030	0	0,049	0,042	0,032	0,029	0,037
Соф.204	0,083	0,080	0,049	0,049	0	0,053	0,048	0,046	0,054
БАН4.5.10	0,082	0,076	0,042	0,042	0,053	0	0,044	0,035	0,047
Соф.382	0,077	0,071	0,032	0,032	0,048	0,044	0	0,027	0,039
БАН16.14.13	0,076	0,070	0,030	0,029	0,046	0,035	0,027	0	0,036
T.112	0,074	0,067	0,037	0,037	0,054	0,047	0,039	0,036	0
Общая дистанция	0,641	0,603	0,371	0,370	0,465	0,426	0,374	0,353	0,395

Таблица 4. Евклидова дистанция между списками (лексика и грамматика)

Наименьшее значение для каждого списка в столбце выделено курсивом. Наименьшее значение в последней строке, показывающую общую дистанцию между списками, выделено курсивом.

В Таблице 5 представлены результаты сравнения списков в случае, когда учитывается лексические и грамматические различия.

Проанализируем лексику и грамматику, опираясь на Таблицы 4 и 5.

1. Наименьшее общее сходство с группой исследуемых текстов имеют Q.п.I.25 и ПМ: отмечено различие в выборе падежной формы  $\overline{гь}$   $\overline{стъ}$  (ПМ) –  $\overline{гьдвѣтъжѣнхъ}$   $\overline{вьсѣхъ}$   $\overline{гь}$  (Q.п.I.25) –  $\overline{гьдѣствоуѣннмъ}$   $\overline{сѣ}$   $\overline{ннмн/ннмь}$  (остальные списки).

2. Следует отметить, что ПМ имеет теперь наибольшее сходство с T.112: по форме аориста  $\overline{спсе}$  (ПМ, T.112) и падежной форме  $\overline{дшъ}$   $\overline{нашнхъ}$  (ПМ, T.112).

3. Далее по сходству с ПМ следует Q.п.I.25: по оформлению звательной формы  $\overline{исѣ}$ , в других списках отмечены формы  $\overline{исѣ}$  –  $\overline{нсѣ}$ .

4. Так же, как и при анализе только лексических различий, выявляется сходство Q.п.I.25 с T.112:  $\overline{исѣ}$  –  $\overline{исѣ/нсѣ}$  (в других списках этого типа).

Наибольшее общее сходство с остальными списками наблюдается у БАН16.14.13. Эта рукопись наиболее близка к Соф.203, Син.166, Соф.204, БАН4.5.10 и Соф.382. Например, в этих списках  $\overline{соломоню}$  /  $\overline{соломоноу}$   $\overline{прмд}$   $\overline{реть}$  приводится без разночтений, ср. в списке Q.п.I.25  $\overline{соломонь}$   $\overline{прѣмьд}$   $\overline{реть}$ .

	ПМ	Q.п.1.25	Соф.203	Син.166	Соф.204	БАН4.5.10	Соф.382	БАН16.14.13	Т.112
ПМ	1	0,145	0,070	0,083	0,096	0,053	0,062	0,078	0,231
Q.п.1.25	0,145	1	0,078	0,075	0,035	0,048	0,070	0,067	0,257
Соф.203	0,070	0,078	1	0,277	0,119	0,153	0,269	0,280	0,226
Син.166	0,083	0,075	0,277	1	0,106	0,169	0,243	0,310	0,244
Соф.204	0,096	0,035	0,119	0,106	1	0,213	0,213	0,244	0,087
БАН4.5.10	0,053	0,048	0,153	0,169	0,213	1	0,135	0,435	0,138
Соф.382	0,062	0,070	0,269	0,243	0,213	0,135	1	0,449	0,201
БАН16.14.13	0,078	0,067	0,280	0,310	0,244	0,435	0,449	1	0,249
Т.112	0,231	0,257	0,226	0,244	0,087	0,138	0,201	0,249	1
Общая схожесть	1,821	1,778	2,475	2,509	2,116	2,348	2,645	3,114	2,637

Таблица 5. Косинус-сходство списков (лексика и грамматика)

Наибольшее значение для каждого списка в столбце выделено курсивом. Наибольшее значение в последней строке, показывающую общую схожесть списков, выделено курсивом.

В Таблице 6 представлены наиболее значимые лексические различия в списках. В векторном представлении списков веса этих терминов максимальны, поэтому эти термины являются особенными в каждом из списков и отделяют его от остальной исследуемой группы. Интересно отметить, что наибольшее значение имеют лексические различия. Только в случае Соф.204 и БАН16.14.13. грамматические различия *правдъноє* – полная форма прилагательного на месте краткой *правдъно*; и *снягъ* – причастие в другой форме – соответственно, становятся наиболее значимыми. В этих списках отражены изменения в грамматической системе языка.

Список	Термины Лексика	Лексика и грамматика
ПМ	прослав'ля ꙗемъ	прослав'ля ꙗемъ
Q.п.1.25	въ   семген	въ   семген
Соф.203	дарочн	дарочн
Син.166	богочьстнкоумъ	богочьстнкоумъ
Соф.204	въснявъ	правдъноє
БАН4.5.10	правдою	правдою
Соф.382	неподнмо	неподнмо
БАН16.14.13	члвколюбное	снягъ
Т.112	кръвьское	кръвьское

Таблица 6. Наиболее значимые различия в списках

В различии *снявъ* (ПМ) – *снъвъ* (Q.п.1.25) – *въснявъ* (Соф.204) – *снягъ* (БАН4.5.10, БАН16.14.13) греческое чтение *λάμψας* (*λάμπω* ‘сиять’) отражено в списках архаичного типа, в студийском списке Соф.204 употребляется приставочный вариант причастия. Семантика приставочного беспри-

ставочного глагола совпадает  $vz\text{zn}atn = vz\text{z} + \text{zn}atn$ . Разночтение отражает грамматические тенденции в глагольной системе: приставка, семантически не нагруженная, становится видообразующей. Отражена также временная вариативность форм действительного причастия: прошедшее время сменяется настоящим в списках БАН4.5.10, БАН16.14.13.

Таким образом из общей группы текстов можно выделить списки ПМ и Q.п.I.25, которые отстоят от других по лексическому и грамматическому сходству. Лексически эти списки схожи между собой. Если учитывать лексику и грамматику, то наиболее схож с этими двумя списками список Т.112. Далее можно выделить в общую группу списки Соф.203, Син.166, Соф.204, БАН4.5.10, Соф.382, БАН16.14.13, Т.112. Отдельно следует отметить список БАН16.14.13, который является наиболее близким к большинству исследуемых списков.

Предыдущие исследования, используя другую методику, приходили к выводу, что списки ПМ и Q.п.I.25 образуют особый тип памятника по структуре текста и лингвистическим чтениям. Списки БАН16.14.13 и Т.112 выделяются как типовые для второго типа Минеи: полное совпадение в наборе памятей по спискам происходит 23 мая преп. Иоанну Психитянину: ПМ, Соф.203, Син.166, Соф.204, БАН4.5.10, Соф.382, БАН16.14.13, Т.112. Математический анализ лексики и грамматики подтверждает схожесть списков БАН16.14.13 и Т.112 со списками первого типа (ПМ). Выделить по всем параметрам в окончательном виде текст-стандарт для второго типа рукописей, в отличие от первого, где эту функцию выполняет ПМ, достаточно сложно, используя только лингвотекстологический метод. Однако математические методы показывают, что тексты списков БАН16.14.13 и Т.112 могут быть представлены как типовые для списков, соотнесенных со Студийским уставом.

##### 5. ЗАКЛЮЧЕНИЕ

В данной статье проанализировано девять списков Минеи с использованием формальной векторной модели. Этими списками являются ПМ, Q.п.I.25, Соф.203, Син.166, Соф.204, БАН4.5.10, Соф.382, БАН16.14.13 и Т.112. Анализ и классификация данных текстов в предшествующей литературе были основаны на лингвотекстологической основе. В данной статье списки представлены векторами в общем векторном пространстве. Использование векторной модели позволяет более глубоко исследовать схожесть и различие данных памятников. Результаты векторного анализа позволяют выделить ПМ и Q.п.I.25 из общего набора текстов. Эти списки схожи не только текстологически, но и лексически. Также выделяются списки БАН16.14.13 и Т.112 как наиболее схожие с большинством анализируемых текстов. Наибольшее значение имеют лексические различия между текстами, тогда как грамматические различия менее показательны. Первые отражают динамику текста: архаичный тип – студийский тип – иерусалимский, а вторые – динамику развития языковой системы.

SLAVIC MENAION MANUSCRIPTS FROM THE 11TH–14TH C. AND PRINCIPLES  
OF THEIR CLASSIFICATION

*(Summary)*

In the present paper we analyze nine Menaia manuscripts (Old-Slavic hymnographic texts) from the 11th–14th c. using a vector space model. The analysis and classification of the manuscripts in previous studies were rather subjective and based on linguo-textological properties of the texts. Our approach uses contemporary information retrieval methods and represents the manuscripts as vectors in a common vector space. A vector space model allows to look at similarities and differences between the manuscripts from a different angle. This approach is more data driven and less subjective than the standard approach. The vector analysis allows to distinguish the Putyatina Menaion and the Menaion Q.п.I.25 from the set of analyzed texts. These manuscripts share both textological and lexical similarities. Similar findings are reached in the existing studies. The manuscripts БАН16.14.13 and T.112 are shown to be rather similar to the set of analyzed texts. This result is new. Most important are lexical differences between the set of texts analyzed while differences in grammar are less pronounced. This is important as the first feature shows dynamics of the texts from the archaic to the Jerusalem type while the latter feature shows the dynamics of the language system.

*Natalya Nechunajeva, Tallinn University,  
Aleksy Nechunajev, Free University of Berlin*