

Simeon Stefanov

(Bulgaria, Institute for Bulgarian Language, BAS)

The Vocabulary of the Life of St. Catherine (Legenda sv. Katarine divice) from Manuscript Ic6 in the HAZU Archive: a Glottometric Analysis

Abstract: This paper is an attempt at applying statistical methods (Фрумкина 1964: 42-54; Янакиев 1977; Давидов 1996: 28-43; Илиева 2011: 14-36; 2013: 72-86) to the study of the vocabulary of one of the earliest known medieval Croatian texts written in the Croatian vernacular with Latin script, Legenda sv. Katarine divice from manuscript I c 6 in the HAZU Archive. The main aim is to provide comparative material for similar future studies.

Keywords: one of the earliest texts in medieval Croatian literature; historical lexicology of the Croatian language; glottometric analysis.

Симеон Стефанов

(България, Институт за български език, БАН)

Лексиката в житието на св. Екатерина *Legenda sv. Katarine divice* по ръкопис I c 6 от Архива на ХАЗУ: Глотометричен анализ

1. Предмет, цели, задачи и методи на изследване

Настоящата статия е опит да се приложи методът на статистическата лингвистика (Фрумкина 1964: 42-54; Янакиев 1977; Христова 1994: 7-40; Давидов 1996: 28-43; Илиева 2011: 14-36; 2013: 72-86 и др.) върху лексикалния материал от житието на св. Екатерина:

LSK *Legendasv. Katarinedivice* по ръкопис I c 6 от Архива на Хърватската академия на науките и изкуствата (ХАЗУ), нататък *LSK* (срв. и Стефанов 2015б).

Тъй като не ни е известно да са правени подобни проучвания в хърватската медиевистика, основната цел на работата е да даде материал за сравнение и съпоставка на получените резултати. Все пак за някои от разгледаните категории ще направя съпоставка с резултатите от работата ми върху други три паметника, писани на народен хърватски език и записани с латински букви. Това са:

RZ *Red i zakon od primljenja na dil dobroga činjenja sestar naših reda svetog otca našega Dominika* от 1345 г. - най-ранният известен засега хърватски текст, писан на народен език и записан с латински букви (срв. Стефанов 2015б);

- CPS* *Cantilena pro sabatho* от края на XIV (срв. Стефанов 2015б), преписана от по-ранна подложка около 1385 г. (срв. Стефанов 2015б) и
- ŠM* *Šibenska molitva*, или *Oracio pulcra (et) devota ad B(ea)tam V(irginem)* от края на XIV в. (срв. Стефанов 2018б).

За да прецизирам обобщенията и изводите си се запознах и с резултатите от изследвания върху езика на произведения от старата българска литература, с известното за състоянието в съвременни български литературни жанрове, с жанровата класификация на средновековната хърватска литература. Сведения за състоянието в текстовете от старата и съвременната българска литература заемам от Илиева 2013 и Янакиев 1977, а за жанровата класификация на средновековната хърватска литература от: Hercigonja 1973; Hercigonja 1975; Fališevac 1978. Трябва да се има пред вид, че разглежданият тук текст вероятно е „доста свободен превод“ от латински или италиански език. Въпреки това е възможно някои от езиковите му особености да са присъщи не само на хърватския превод, но и на (латинската/италианската) подложка. Тъй като съпоставката с протограф не е част от задачите на настоящото изследване, а и понеже хърватският превод е направен на народен хърватски език, смятам, че *LSK*, който отразява до голяма степен особеностите на говоримия език от разглеждания период, може да ни послужи и като свидетелство и за неговото състояние.

2. За паметника. Лексиката в *LSK*. Глотометричен анализ

2.1. За паметника

LSK е средновековен текст на житието на св. Екатерина, александрийската девица и мъченица, писан на народен хърватски език. Запазен само в един препис от XV в., но се предполага, че е възникнал по-рано (Jagić 1869: 217). Днес се съхранява в Архива на ХАЗУ (Хърватската академия на науките и изкуствата) под сигнатура I с б. Диалектно текстът може да бъде причислен към Средишната чакавска икавско-екавска област от периода XIV-XV в. (вж. Lukežić 2012: 225-252). Ръкописът е открит в библиотеката на девическия бенедиктински манастир „Св. Мария“ в Задар, където вероятно е бил занесен от монахините на манастира „Св. Екатерина“ в Задар, след закриването му.¹ По-късно текстът е станал част от богатата книжна колекция на Иван Кукулевич, а от 1869 г. се пази сред ръкописите на библиотеката на Академията на науките и изкуствата, първоначално под номер 317 (Jagić 1869: 217), а след това под сигнатура Ic6. Публикуван е три пъти до сега (Jagić 1869: 217-224; Štefanić 1969: 274-277; Vadurina-Stipčević 2013: 106-114). Предполага се, че е компилиран от различни латински образци (Hamm 1926/1927: 16-17; Fališevac 2010: 257-258). Не бива да се

¹ В Задаре имало пет женски манастира. Това са: бенедиктинският „Св. Мария“, францисканският „Св. Никола“, доминиканският „Св. Димитър“, манастирът „Св. Марчела“ на францисканските клариси и бенедиктинският „Св. Екатерина“. Последните два са били предназначени за девойки от простолудието. Останалите три - предимно за девойки от изтъкнати задарски семейства. Доминиканският манастир „Св. Димитър“ е основан през XII в. и първоначално е бил притежание на бенедиктинките, които през 1125 г., след разрушаването на Биоград, се преместват в Задар. От 1311 г. той е доминикански. Закрит е на 18.III.1779 г. Францисканският манастир „Св. Никола“ е закрит през 1789 г. Бенедиктинският манастир „Св. Екатерина“ – през 1804 г., но монахините му са прехвърлени в дотогава само благородническият бенедиктински манастир „Св. Мария“. Днес съществува само манастирът „Св. Мария“, основан през XI в. от абатката Чика. В началото той е предназначен за обучението само на девойки от благороднически произход. По-късно – и от други изтъкнати задарски семейства. Още от основаването си манастирът се е превърнал в място за обучение и възпитание както на момичета, които се подготвят да полагат изпит за приемане в ордена, така и за онези, които не възнамеряват да се замонашват. Основен език в него е бил хърватският (Ganza-Aras 1990: 198-208).

изключва и вероятността за наличие на италианска подложка. Моето предположение е, че житието „*Legenda sv. Katarine divice*” е написано на език, употребяван в книжнината от областта Сплит-Задар-Шибеник в периода XIV-XV в. (Стефанов 2014: 43; 2015а: 51-52), и затова е важно свидетелство за състоянието на хърватския език от този период.

2.2. Лексиката в LSK. Глотометричен анализ

2.2.1. Показатели на словника

Словникът на *LSK* наброява 626 заглавни единици в 2155 употреби. От тях по честота на употреба в текста в групата от първата до петдесетата дума влизат 15 съществителни имена, 5 прилагателни имена, 12 глаголни форми, или общо 38 лексеми пълнозначна лексика², срещу 18 - служебна лексика. Изчислени в проценти тези стойности показват, че 30% от петдесетте най-често употребявани думи в текста са съществителни имена. С общо 152 словоупотребии те заемат 7,05% от текста, а като 15 съществителни имена – 2,4% от целия словник. Прилагателните имена представят 10% от групата. С общо 42 словоупотребии те заемат 1,95% от текста, а като 5 прилагателни имена – 0,8% от целия словник. Глаголните форми са 12 и съставляват 24% от групата. С общо 264 словоупотребии те заемат 12,25% от текста, а като 12 заглавни единици – 1,92% от целия словник. Сред петдесетте най-често употребявани думи в текста пълнозначната лексика наброява 38 заглавни единици, или 76% от групата. С 458 словоупотребии тя заема 21,25% от текста, а като 38 заглавни единици – 6,07% от целия словник. Непълнозначната лексика съставлява 36% от групата. Със 711 словоупотребии тя заема 32,99% от текста, а като 18 заглавни единици представя 2,88% от целия словник.

В групата от 51 до 100 като най-чести лексеми в словника на *LSK* се явяват 23 съществителни имена, 8 прилагателни имена, 9 глаголни форми, или общо 40 лексеми пълнозначна и 10 – непълнозначна лексика. Изчислени в проценти тези стойности показват, че 46% от думите в тази група са съществителни имена. С общо 104 словоупотребии те заемат 4,83% от текста, а като 23 съществителни имена - 3,67% от целия словник. Прилагателните имена съставят 16% от групата. С общо 28 словоупотребии те заемат 1,3% от текста, а като 8 заглавни единици – 1,28% от целия словник. Глаголните форми са 18% от групата. С общо 110 словоупотребии те заемат 5,1% от текста, а като 9 заглавни единици – 1,84% от целия словник. Пълнозначната лексика в групата е представена от 40 лексеми, или 80% от нея. С общо 242 словоупотребии тя заема 11,23% от текста, а като 40 заглавни единици – 6,39% от целия словник. Непълнозначната лексика в групата е 10 лексеми и съставлява 20% от нея. С общо 172 словоупотребии тя заема 7,98% от текста, а като 10 заглавни единици - 1,6% от целия словник.

В групата от 101 до 150 по честота на употреба единица в словника влизат 15 съществителни имена, 5 прилагателни имена, 16 глаголни форми, или общо 36 единици пълнозначна лексика и 14- непълнозначна лексика. Изчислени в проценти тези стойности показват, че 30% от тази група са съществителни имена. С общо 36 словоупотребии те заемат 1,67% от текста, а като 15 заглавни единици - 2,4% от целия словник. Прилагателните имена са 10% от групата. С общо 14 словоупотребии те заемат 0,65% от текста, а като 5 заглавни единици – 0,8% от целия словник. Глаголните форми са 32% от групата. С общо 42 словоупотребии те заемат 1,95% от текста, а като 16

² В пълнозначната лексика в разглеждания текст влизат съществителни имена, прилагателни имена и глаголи, а в служебната - местоимения, местоименни наречия, съюзи, предлози, частици, междуметия.

заглавни единици – 2,56% от целия словник. Пълнозначната лексика е 72% от групата. С общо 92 словоупотреби тя заема 4,27% от текста, а като 36 заглавни единици – 5,75% от целия словник. Непълнозначната лексика е 28% от групата. С общо 40 словоупотреби тя заема 1,86% от текста, а като 14 заглавни единици – 2,64% от целия словник.

В групата от 151 до 214 по честота на употреба единица в словника влизат 22 съществителни имена, 7 прилагателни имена, 19 глаголни форми, или общо 48 единици пълнозначна лексика и 16 - непълнозначна лексика. Изчислени в проценти тези стойности показват, че 34,38% от тази група са съществителни имена. С общо 44 словоупотреби те заемат 2,64% от текста, а като 22 заглавни единици - 3,51% от целия словник. Прилагателните имена са 10,94% от групата. С общо 14 словоупотреби те заемат 0,65% от текста, а като 7 заглавни единици – 1,12% от целия словник. Глаголните форми са 29,69% от групата. С общо 38 словоупотреби те заемат 1,76% от текста, а като 19 заглавни единици – 3,04% от целия словник. Пълнозначната лексика е 75% от групата. С общо 96 словоупотреби тя заема 4,45% от текста, а като 48 заглавни единици – 7,67% от целия словник. Непълнозначната лексика е 25% от групата. С общо 32 словоупотреби тя заема 1,48% от текста, а като 16 заглавни единици – 2,56% от целия словник.

В групата от 215 по честота дума до края на словника, която наброява 411 заглавни единици, влизат 143 съществителни имена, 71 прилагателни имена, 145 глаголни форми, 53 единици служебна лексика. Това са лексеми употребени само по веднъж в текста. Изчислени в проценти тези стойности показват, че 34,71% от тази група са съществителни имена. С общо 143 словоупотреби те заемат 6,64% от текста, а като 143 заглавни единици- 22,84% от целия словник. Прилагателните имена са 17,96% от групата. С общо 71 словоупотреби те заемат 3,29% от текста, а като 71 заглавни единици – 11,34% от целия словник. Глаголните форми са 35,19% от групата. С общо 145 словоупотреби те заемат 6,73% от текста, а като 145 заглавни единици – 23,16% от целия словник. Общият брой на пълнозначната лексика в тази група е 369 лексеми, или 89,78% от нея. Със 369 словоупотреби те заемат 17,12% от текста, а като 369 заглавни единици – 58,95% от целия словник. Непълнозначната лексика в групата е 53 лексеми, или 12,86% от нея. С общо 53 словоупотреби те заемат 2,46% от текста, а като 53 заглавни единици - 8,47% от целия словник.

Общото количество на съществителните имена в словника е 218 заглавни единици, или 34,82% от него, на прилагателните имена – 96 заглавни единици, или 15,34% от него, на глаголните форми – 201 заглавни единици, или 32,11% от него. Общото количество на пълнозначната лексика е 515 заглавни единици, или 82,27% от него, а на непълнозначната – 111 заглавни единици, или 17,73% от него.

Въз основа на изнесените резултати, въпреки малкия обем на паметника, може да се забележи тенденцията да е висок процентът на служебната лексика сред най-често употребяваните думи в текста. Колкото повече намалява честотата на употребените единици в текста, намалява и количеството на служебната лексика, а нараства това на пълнозначната. Относително ниският процент на служебната лексика в словника показва, че голяма част от единиците от тази категория са употребени по повече от веднъж в текста.

Отношението *покритие в текста - покритие в словника* показва, че колкото по-висока е честотата на употреба на лексемите, по-голямо е и покритието в текста, а по-

малко - в словника. С намаляването на честотата на употреба, намалява покритието на групата в текста, а нараства покритието в словника.

2.2.2. Показатели на отношението между словника и текста

2.2.2.1. Коефициент на лексикално многообразие

Важна категория при глотометричния анализ на един текст е и т.н. коефициент на лексикално многообразие (Илиева 2011: 15; 2013: 73 и др.) Чрез него се оценява съотношението между употребените поне веднъж лексеми и всички употреби в текста. Той се изчислява като разделим броя лексеми в словника на общия брой употреби и умножим полученото число по 100, за да го превърнем в проценти. По този начин се игнорират повторените единици, а се взимат пред вид само първите употреби на всяка лексема. В разглеждания текст коефициентът на лексикално многообразие е 29,05%. Това е по-ниска стойност в сравнение с другите три разгледани текста. Тя може да е повлияна и от по-големия обем на паметника. Възможно е, все пак, тази категория да се влияе и от жанровите особености на текстовете. За сведение коефициентът на лексикално многообразие в *RZ* е 40,06 %, в *ŠMe* 32,97%, а в *CPSe* 48,51%.

Приведените резултати не показват особена стабилност за разглежданата категория в най-ранните хърватски текстове, писани на народен език и записани с латински букви, като е възможно разликата да се дължи на различния обем на паметниците, а вероятно също и на жанровите различия между тях. Прави впечатление, че стойността на разглежданата категория е най-висока в най-краткия текст, който се придържа по-строго към границите на жанра, а най-ниска – в най-обемния, в който се наблюдава съчетаването на откъси, характерни за различни средновековни жанрове. Общото между резултатите е, че не надхвърлят 50%.

Приведените резултати все пак са свидетелство за това, че въпреки че не се опират на предходна книжовна традиция от подобен характер, хърватските книжовници успяват да създадат текстове, които дори и в повечето случаи преводни, да отразяват особеностите и на говоримия език, но и да са относително лексикално многообразни. Все пак това са различни по обем текстове, така че изнесените данни биха могли да служат само като ориентир и отправна точка при бъдещи съпоставки на равни части от тези и други подобни произведения.

2.2.2.2. Индекс на повтаряемостта

Друга съществена категория при глотометричния анализ на едно произведение е т.н. индекс на повтаряемостта (Давидов 1996: 38; Илиева 2013: 76; Фрумкина 1964: 42-43; 50-54 и др.) Той се пресмята като разделим общия брой словоупотреби на броя единици в словника. Индексът на повтаряемостта в разглежданото произведение е 3,44. За сведение в *RZ* той е 2,5, в *ŠM* 3,03, а в *CPSe* 2,06.

Това са относително ниски стойности на разглежданата категория в средновековната хърватска литература, писана на народен език с латински букви, което показва, че вероятно тя е относително стабилна в текстовете на народен език от този период. Все пак при някои от произведенията ниските стойности може да се дължат и на малкия обем на разглеждания паметник.³

³ В съвременната българска поезия, където този индекс е най-малък, той се движи около 3; в журналистиката – 4,1, а в научната литература – 6,2 (Давидов 1996: 38; Илиева 2013: 76). В съчиненията от старобългарската литература този показател е: в съчиненията на Петър Черноризец – 5,15; в

И тук трябва да уточним, че това са различни по обем текстове, така че приведените данни биха могли да служат само като ориентир и отправна точка при бъдещи съпоставки на равни части от тези и други подобни произведения. Напълно очаквано е колкото по-ниска е стойността на индекса на повторяемост, толкова по-голямо да е разнообразието на речника на дадено произведение.

2.2.3. Показатели на текста

2.2.3.1. Статистически изчисления

С понятието „глаголна температура” се означава процентът на лични глаголни форми, употребени в даден текст или текстови отрязък. Терминът е въведен от Мирослав Янакиев (1977), който уточнява, че по този начин студентите му назовават споменатата текстова категория (срв. Янакиев 1977: 155, 185-190). По модела на този термин бихме могли да въведем (и изчислим) и „номиналната температура”, или процента на съществителните имена, „адективната температура”, или процент на прилагателните имена. Бихме могли да разграничим и процента на личните от този на неличните глаголни форми в общото количество единици от тази категория в даден текст или текстови отрязък. Въз основа на изследванията си върху стила на съвременни български автори, Мирослав Янакиев установява, че текст или текстови отрязък с относително висок процент лични глаголни форми (с висока „глаголна температура”) звучи по-живо, по-динамично, по-раздвижено, по-действено, а насищането на разказа със съществителни имена внася предметност (Янакиев 1977: 155). За относително висок процент лични глаголни форми се смята резултат в рамките 18-25% в една килолекса текст (Янакиев 1977: 185-190). В цитираното изследване Мирослав Янакиев уточнява, че в разговорната практика употребяваме повече лични, отколкото нелични глаголни форми, докато по-честата употреба на нелични глаголни форми е присъща на писаните текстове (Янакиев 1977: 181-185). Въпреки че това заключение е направено върху езика и стила на съвременни български автори и понеже не ми е известен подобен труд в хърватската наука за литературата, обхващаща текстове от периода XIV-XV в., цитирам българския изследовател. Смятам, че наблюденията му биха могли да се окажат валидни и за средновековни (хърватски) текстове, писани на народен език.

Словникът на *LSK* наброява 261 заглавни единици в 2155 словоупотреби. Общото количество на съществителните имена в текста е 498 единици, или 23,11% от него, на прилагателните имена – 172 единици, или 7,98% от него, на глаголните форми – 504 единици, или 23,39% от него. Общото количество на пълнозначната лексика в текста е 1174 единици, или 54,48% от него, а на непълнозначната – 981 единици, или 45,52% от него. От приведените резултати се вижда, че процентът на глаголните форми 23,39% е относително висок. За сведение в *RZ* той е 18,18%, в *ŠM* е 14,33%, а в *CPSe* е 27,7%. Процентът на съществителните имена в разглеждания текст е 23,11%. За сведение в *RZ* той е 30,97%, в *ŠMe* 33,59%, а в *CPSe* 26,58%. Процентът на прилагателните имена в разглеждания текст е 7,98%. За сведение в *RZ* той е 10,51%, в *ŠM* е 18,49%, а в *CPSe* 7,25%.

химнографията на Климент Охридски – 6,9; в словата на Климент Охридски е – 7,07; в творчеството на Презвитер Козма – 7,48; в творчеството на Черноризец Храбър – 4,07; в Повесть временных лет – 10 (срв. Илиева 2013: 76). Причината за толкова висок индекс на повторяемост изобщо в произведенията от старата (българска) литература е „честата употреба на служебни думи и абстрактна лексика за фундаментални мирогледни понятия” (Давидов 1996: 38; Илиева 2013: 76).

Тук отново ще припомня, че това са различни по обем текстове, така че приведените данни биха могли да служат само като ориентир и отправна точка за бъдещи съпоставки на равни части от тези и други подобни произведения. Все пак напълно е възможно някои от изчислените категории да се влияят и от жанровите специфики на произведенията, затова и реших да ги приведа.

2.2.3.2. Средна аритметична стойност и средно квадратично отклонение от нея за разглежданите категории в текста. Гаранция за 100% надеждност на ексцерптна оценка на средната аритметична стойност⁴

Да установим какво е състоянието на разглежданите категории в *LSK* и по килолекси, и за целия текст. Тъй като обемът на паметника е 2155 единици, ще приведа информацията за двадесет и една цели килолекси, но ще изчисля и използвам също и тази седем групи от по три, четири групи от по пет и три групи от по шест цели килолекси, понеже идеята ми е след това да съпоставя получените резултати с резултатите от *RZ*, а неговият обем е 349 единици, или три цели килолекси и половина, от *ŠM*, а неговият обем е 649 единици, или три цели килолекси и половина и от *CPS*, а неговият обем е 538 единици, или пет цели килолекси. В бъдеще тези резултати ще могат да послужат за сравнения и съпоставки с подобни по обем ексцерпти от други средновековни хърватски текстове, писани на народен език и записани с латински букви.

В първата килолекса *LSK* са употребени 27 съществителни имена, или 27% от ексцерпта, 14 прилагателни имена, или 14% от ексцерпта, 29 глаголни форми, или 29% от ексцерпта, общо 70 думи пълнозначна лексика, или 70% от ексцерпта и 30 непълнозначна лексика, или 30% от ексцерпта.

Във втората килолекса се срещат 30 съществителни имена, или 30% от ексцерпта, 11 прилагателни имена, или 11% от ексцерпта, 14 глаголни форми, или 14% от ексцерпта, общо 55 думи пълнозначна лексика, или 55% от ексцерпта и 45 непълнозначна лексика, или 45% от ексцерпта.

В третата килолекса се срещат 30 съществителни имена, или 30% от ексцерпта, 9 прилагателни имена, или 9% от ексцерпта, 18 глаголни форми, или 18% от ексцерпта, общо 57 единици пълнозначна лексика, или 57% от ексцерпта и 43 единици непълнозначна лексика, или 43% от ексцерпта.

В четвъртата килолекса са употребени 22 съществителни имена, или 22% от ексцерпта, 10 прилагателни имена, или 10% от ексцерпта, 23 глаголни форми, или 23% от ексцерпта, общо 55 думи пълнозначна лексика, или 55% от ексцерпта и 45 непълнозначна лексика, или 45% от ексцерпта.

В петата килолекса се срещат 23 съществителни имена, или 23% от ексцерпта, 5 прилагателни имена, или 5% от ексцерпта, 26 глаголни форми, или 26% от ексцерпта, общо 54 думи пълнозначна лексика, или 54% от ексцерпта и 46 непълнозначна лексика, или 46% от ексцерпта.

В шестата килолекса се срещат 30 съществителни имена, или 30% от ексцерпта, 6 прилагателни имена, или 6% от ексцерпта, 19 глаголни форми, или 19% от ексцерпта,

⁴ За изчислението на тези характеристики, вж. Янакиев 1977: 191-200.

общо 55 думи пълнозначна лексика, или 55% от ексцерпта и 45 непълнозначна лексика, или 45% от ексцерпта.

В седмата килолекса се срещат 24 съществителни имена, или 24% от ексцерпта, 4 прилагателни имена, или 4% от ексцерпта, 29 глаголни форми, или 29% от ексцерпта, общо 57 думи пълнозначна лексика, или 57% от ексцерпта и 43 непълнозначна лексика, или 43% от ексцерпта.

В осмата килолекса се срещат 17 съществителни имена, или 17% от ексцерпта, 5 прилагателни имена, или 5% от ексцерпта, 30 глаголни форми, или 30% от ексцерпта, общо 52 думи пълнозначна лексика, или 52% от ексцерпта и 48 непълнозначна лексика, или 48% от ексцерпта.

В деветата килолекса се срещат 21 съществителни имена, или 21% от ексцерпта, 3 прилагателни имена, или 3% от ексцерпта, 24 глаголни форми, или 24% от ексцерпта, общо 48 думи пълнозначна лексика, или 48% от ексцерпта и 52 непълнозначна лексика, или 52% от ексцерпта.

В десетата килолекса се срещат 14 съществителни имена, или 14% от ексцерпта, 9 прилагателни имена, или 9% от ексцерпта, 24 глаголни форми, или 24% от ексцерпта, общо 47 думи пълнозначна лексика, или 47% от ексцерпта и 53 непълнозначна лексика, или 53% от ексцерпта.

В единадесетата килолекса се срещат 13 съществителни имена, или 13% от ексцерпта, 7 прилагателни имена, или 7% от ексцерпта, 29 глаголни форми, или 29% от ексцерпта, общо 49 думи пълнозначна лексика, или 49% от ексцерпта и 51 непълнозначна лексика, или 51% от ексцерпта.

В дванадесетата килолекса се срещат 20 съществителни имена, или 20% от ексцерпта, 12 прилагателни имена, или 12% от ексцерпта, 23 глаголни форми, или 23% от ексцерпта, общо 55 думи пълнозначна лексика, или 55% от ексцерпта и 45 непълнозначна лексика, или 45% от ексцерпта.

В тринадесетата килолекса се срещат 13 съществителни имена, или 13% от ексцерпта, 9 прилагателни имена, или 9% от ексцерпта, 28 глаголни форми, или 28% от ексцерпта, общо 50 думи пълнозначна лексика, или 50% от ексцерпта и 50 непълнозначна лексика, или 50% от ексцерпта.

В четиринадесетата килолекса се срещат 25 съществителни имена, или 25% от ексцерпта, 9 прилагателни имена, или 9% от ексцерпта, 19 глаголни форми, или 19% от ексцерпта, общо 53 думи пълнозначна лексика, или 53% от ексцерпта и 47 непълнозначна лексика, или 47% от ексцерпта.

В петнадесетата килолекса се срещат 20 съществителни имена, или 20% от ексцерпта, 8 прилагателни имена, или 8% от ексцерпта, 24 глаголни форми, или 24% от ексцерпта, общо 52 думи пълнозначна лексика, или 52% от ексцерпта и 48 непълнозначна лексика, или 48% от ексцерпта.

В шестнадесетата килолекса се срещат 24 съществителни имена, или 24% от ексцерпта, 9 прилагателни имена, или 9% от ексцерпта, 24 глаголни форми, или 24% от ексцерпта, общо 57 думи пълнозначна лексика, или 57% от ексцерпта и 43 непълнозначна лексика, или 43% от ексцерпта.

В седемнадесетата килолекса се срещат 27 съществителни имена, или 27% от ексцерпта, 8 прилагателни имена, или 8% от ексцерпта, 23 глаголни форми, или 23% от ексцерпта, общо 58 думи пълнозначна лексика, или 58% от ексцерпта и 42 непълнозначна лексика, или 42% от ексцерпта.

В осемнадесетата килолекса се срещат 23 съществителни имена, или 23% от ексцерпта, 6 прилагателни имена, или 6% от ексцерпта, 23 глаголни форми, или 23% от ексцерпта, общо 52 думи пълнозначна лексика, или 52% от ексцерпта и 48 непълнозначна лексика, или 48% от ексцерпта.

В деветнадесетата килолекса се срещат 23 съществителни имена, или 23% от ексцерпта, 13 прилагателни имена, или 13% от ексцерпта, 22 глаголни форми, или 22% от ексцерпта, общо 58 думи пълнозначна лексика, или 58% от ексцерпта и 42 непълнозначна лексика, или 42% от ексцерпта.

В двадесет и първата килолекса се срещат 23 съществителни имена, или 23% от ексцерпта, 5 прилагателни имена, или 5% от ексцерпта, 23 глаголни форми, или 23% от ексцерпта, общо 51 думи пълнозначна лексика, или 51% от ексцерпта и 49 непълнозначна лексика, или 49% от ексцерпта.

В двадесетата килолекса се срещат 30 съществителни имена, или 30% от ексцерпта, 7 прилагателни имена, или 7% от ексцерпта, 26 глаголни форми, или 26% от ексцерпта, общо 63 думи пълнозначна лексика, или 63% от ексцерпта и 37 непълнозначна лексика, или 37% от ексцерпта.

В останалите 55 единици текст се срещат 19 съществителни имена, или 34,55% от ексцерпта, 3 прилагателни имена, или 5,45% от ексцерпта, 6 глаголни форми, или 10,9% от ексцерпта. Общо 28 единици пълнозначна лексика, или 50,9% от ексцерпта и 27 непълнозначна лексика, или 49,1% от ексцерпта.

Броят на глаголните форми в двадесет и една цели килолекси *LSK* е съответно 29, 14, 18, 23, 26, 19, 30, 30, 24, 24, 29, 23, 28, 19, 24, 24, 21, 23, 22, 23, 26. Броят на съществителните имена е 27, 30, 30, 22, 23, 30, 24, 17, 21, 14, 13, 20, 13, 25, 20, 24, 27, 23, 23, 23, 30, а броят на прилагателните имена – 14, 11, 9, 10, 5, 6, 4, 5, 3, 9, 7, 12, 9, 9, 8, 9, 8, 6, 13, 5, 7. Средното аритметично количество глаголни форми за двадесет и една килолекси текст е 25,73. Отгук като извадим от средното аритметично реалното количество глаголни форми за всяка килолекса, получените резултати за всеки един ексцерпт вдигнем на квадрат и съберем стойностите помежду им, след това разделим полученото число на двадесет, понеже работим с двадесет и една килолекси, а на края сложим под квадратен корен резултата, получаваме 12,7. Това е средното квадратично отклонение на глаголните форми за всеки ексцерпт в двадесет и една цели килолекси текст. То ни показва с колко в една килолекса текст техният брой се отклонява от средната стойност за тази категория в *n* на брой (в нашия случай 21) килолекси текст. С негова помощ бихме могли да изчислим интервала на средната стойност на употребените глаголни форми в двадесет и една килолекси *LSK*. В нашия случай за глаголните форми той би трябвало да е 13,03-38,43. Отново с помощта на средното квадратично отклонение можем да изчислим гаранцията за 100% надеждност на ексцерптна оценка на стойността на тази категория. Тя се получава като умножим по три средното квадратично отклонение и за глаголните форми в *LSK*: $12,7 \times 3 = 38,11$.

Средното аритметично на употребените съществителни имена в двадесет и една килолекси текст от *LSK* е 22,81. Средното квадратично отклонение за тази категория в

текстовите ексцерпти е 5,41. Интервалът на средната стойност на употребените съществителни имена за двадесет и една килолекси текст *LSK* е 17,4-28,22. Гаранцията за 100% надеждност на ексцерптна оценка на стойността на съществителните имена за *LSK* е $5,41 \times 3 = 16,22$.

Средното аритметично за прилагателните имена в двадесет и една килолекси *LSK* е 8,05. Средното квадратично отклонение за тази категория е 2,94. Интервалът на средната стойност на употребените прилагателни имена е 5,11-10,99. Гаранцията за 100% надеждност на ексцерптна оценка на стойността на прилагателните имена за *LSK* е $2,94 \times 3 = 8,8$.

Приведените резултати ще ми послужат по-късно за да изчисля по „Критерия на Пиърсън“⁵ колко единици от всяка една категория бихме могли да очакваме в останалите 45 единици текст от *LSK* на последната двестет и втора килолекса.

Заради съпоставката с резултатите от работата ми върху *RZ*, ще направя същите изчисления и за само три килолекси текст, понеже *RZ* има само три цели килолекси текст. След това ще приведа резултатите от изчисленията на тези категории в по пет и в по шест килолекси, понеже *CPS* има пет цели, *SM*, а шест цели килолекси текст.

2.2.3.2.1. Средна аритметична стойност и средно квадратично отклонение от нея за разглежданите категории в по три килолекси текст

Средното аритметично на употребените съществителни имена в първите три килолекси текст от *LSK* е 29. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,73.

Средното аритметично на употребените прилагателни имена в първите три килолекси текст от *LSK* е 11,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 2,52.

Средното аритметично на употребените глаголни форми в първите три килолекси текст от *LSK* е 20,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 7,77.

Средното аритметично на употребените съществителни имена във вторите три килолекси текст от *LSK* е 28,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 5,97.

Средното аритметично на употребените прилагателни имена във вторите три килолекси текст от *LSK* е 7. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 2,65.

Средното аритметично на употребените глаголни форми във във вторите три килолекси текст от *LSK* е 22,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,51.

Средното аритметично на употребените съществителни имена в третите три килолекси текст от *LSK* е 20,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,51.

⁵ В статистическата лингвистика тази процедура се използва за оценка на отношението между теорията и практиката при изчисляване на дадени стойности, вж. Янакиев 1977: 220-222.

Средното аритметично на употребените прилагателни имена в третите три килолекси текст от *LSK* е 4. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1.

Средното аритметично на употребените глаголни форми в третите три килолекси текст от *LSK* е 27,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,21.

Средното аритметично на употребените съществителни имена в четвъртите три килолекси текст от *LSK* е 15,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,79.

Средното аритметично на употребените прилагателни имена в четвъртите три килолекси текст от *LSK* е 9,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 2,52.

Средното аритметично на употребените глаголни форми в четвъртите три килолекси текст от *LSK* е 25,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,22.

Средното аритметично на употребените съществителни имена в петите три килолекси текст от *LSK* е 19,33. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 6,03.

Средното аритметично на употребените прилагателни имена в петите три килолекси текст от *LSK* е 8,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 0,58.

Средното аритметично на употребените глаголни форми в петите три килолекси текст от *LSK* е 23,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 4,51.

Средното аритметично на употребените съществителни имена в шестите три килолекси текст от *LSK* е 24,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 2,08.

Средното аритметично на употребените прилагателни имена в шестите три килолекси текст от *LSK* е 7,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,53.

Средното аритметично на употребените глаголни форми в шестите три килолекси текст от *LSK* е 22,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,53.

2.2.3.2.2. Средна аритметична стойност и средно квадратично отклонение от нея за разглежданите категории в по пет килолекси текст

Средното аритметично на употребените съществителни имена в първите пет килолекси текст от *LSK* е 26,4. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,78.

Средното аритметично на употребените прилагателни имена във първите пет килолекси текст от *LSK* е 9,8. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,27.

Средното аритметично на употребените глаголни форми в първите пет килолекси текст от *LSK* е 22. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 6,04.

Средното аритметично на употребените съществителни имена във вторите пет килолекси текст от *LSK* е 21,2. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 6,22.

Средното аритметично на употребените прилагателни имена във вторите пет килолекси текст от *LSK* е 5,4. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 2,3.

Средното аритметично на употребените глаголни форми във вторите пет килолекси текст от *LSK* е 25,2. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 4,44.

Средното аритметично на употребените съществителни имена в третите пет килолекси текст от *LSK* е 18,2. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 5,17.

Средното аритметично на употребените прилагателни имена в третите пет килолекси текст от *LSK* е 9. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,87.

Средното аритметично на употребените глаголни форми в третите пет килолекси текст от *LSK* е 24,6. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 4,04.

Средното аритметично на употребените съществителни имена в четвъртите пет килолекси текст от *LSK* е 24. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,73.

Средното аритметично на употребените прилагателни имена в четвъртите пет килолекси текст от *LSK* е 8,2. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,11.

Средното аритметично на употребените глаголни форми в четвъртите пет килолекси текст от *LSK* е 22,6. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,14.

2.2.3.2.3. Средна аритметична стойност и средно квадратично отклонение от нея за разглежданите категории в по шест килолекси текст

Средното аритметично на употребените съществителни имена в първите шест килолекси текст от *LSK* е 27. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,69.

Средното аритметично на употребените прилагателни имена в първите шест килолекси текст от *LSK* е 9,17. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,45.

Средното аритметично на употребените глаголни форми в първите шест килолекси текст от *LSK* е 21,5. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 5,54.

Средното аритметично на употребените съществителни имена във вторите шест килолекси текст от *LSKe* 18,17. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 4,26.

Средното аритметично на употребените прилагателни имена във вторите шест килолекси текст от *LSK* е 6,67. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,46.

Средното аритметично на употребените глаголни форми във вторите шест килолекси текст от *LSKe* 26,5. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,15.

Средното аритметично на употребените съществителни имена в третите шест килолекси текст от *LSKe* 22. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 4,98.

Средното аритметично на употребените прилагателни имена в третите шест килолекси текст от *LSK* е 8,17. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 1,17.

Средното аритметично на употребените глаголни форми в третите шест килолекси текст от *LSKe* 23,17. Средното квадратично отклонение за тази категория в текстовите ексцерпти е 3,06.

2.2.3.3. Критерий на Пиърсън

Сега вече бихме могли да изчислим и потенциално колко съществителни имена, прилагателни имена, глаголни форми, пълнозначна и непълнозначна лексика бихме могли да очакваме в останалите 45 единици текст *LSK* от последната двадесет и втора килолекса текст. Така ще получим още една хипотетична цяла килолекса. Текстовата корекция до края на текста е 45 единици. За да получим търсеното от нас количество на употребени единици в текстовата корекция ще разделим гаранцията за 100% надеждност на изчисление на стойността съответната категория за целия текст на 45 и ще умножим по гаранцията за 100% надеждност на изчисление на съответната категория. След направените изчисления получаваме, че очакваното количество съществителни имена в текстовата корекция е $(16,22/45) \times 16,22 = 5,85 = 6$. Очакваното количество прилагателни имена в текстовата корекция е $(8,8/45) \times 8,8 = 1,72 = 2$. Очакваното количество глаголни форми е $(38,11/45) \times 38,11 = 32,27 = 32$.

Общото количество съществителни имена в хипотетична цяла двадесет и втора килолекса *LSK* е $19+6=25$ (25%). Общото количество прилагателни имена в хипотетична цяла двадесет и втора килолекса *LSKe* $3+2=5$ (5%). Общото количество глаголни форми в хипотетична цяла двадесет и втора килолекса *LSK* е $6+32=38$ (38%). Общото количество на пълнозначна лексика в хипотетична цяла двадесет и втора килолекса *LSKe* $25+5+38=68$ (68%), а на непълнозначната - $27+5=32$ (32%).

2.2.3.4. Критерий на Колмогоров-Смирнов⁶

В статистическата лингвистика тази процедура се използва за установяване на отношението между стойностите на избрани категории в два вариационни реда (от два текста) от по равен брой килолекси (значи при съпоставка на равни по обем ексцерпти). Тук аз ще я приложа, за да установя отношението два по два вариационни реда между

⁶ По-подробно за Критерия Колмогоров-Смирнов, вж. Янакиев 1977: 202-206.

ексцерпти от по три (за *LSK* и *RZ*), по пет (за *LSK* и *CPS*) и по шест (за *LSK* и *ŠM*) килолекси глаголна, номинална и адективна температура от разглежданите текстове.

С $\Delta max.$ се отбелязва максималната разлика между стойностите за съответната категория по килолекси в двата съпоставяни вариационни реда. С помощта на λ^2 можем да заключим съществена ли е или не разликата между средните стойности за дадена категория в два вариационни реда от по равен брой килолекси. Изчислява се като разделим $\Delta^2 max.$ на броя килолекси в редицата, умножени по две, или $\lambda = \Delta^2 max. / 2n$. В лингвостатистиката е установено, че ако λ е по-голямо от 1,84, то е възможно двата съпоставяни ексцерпта да „са от различна генерална съвкупност“, а ако е по-голямо и от 2,65, то двата съпоставяни ексцерпта със сигурност „са от различна генерална съвкупност“, или казано по друг начин, принадлежат на различни по тип текстове.

Тъй като *RZ* ни дава три цели килолекси текст, *CPS* – пет, *ŠM* – шест, а *LSK* двадесет и една, ще съпоставя първо трите килолекси *RZ* с шест пъти по три килолекси *LSK*, след това петте килолекси текст *CPS* с четири пъти по пет килолекси *LSK*, а накрая шестте килолекси текст *ŠM* с три пъти по шест килолекси *LSK*.

2.2.3.4.1. Критерий на Колмогоров-Смирнов за *LSK* и *RZ*

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (първите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 4=9/8=1,125 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адективни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (първите три), $\Delta max.=2$, а $\lambda=2^2/2 \times 5=4/10=0,4 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (първите три), $\Delta max.=1$, а $\lambda=1^2/2 \times 6=1/12=0,083 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (вторите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 5=9/10=0,9 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адективни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (вторите три), $\Delta max.=1$, а $\lambda=1^2/2 \times 5=1/10=0,1 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (вторите три), $\Delta max.=2$, а $\lambda=2^2/2 \times 6=4/12=0,33 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (третите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 5=9/10=0,9 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адективни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (третите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 6=9/12=0,75 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (третите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 6=9/12=0,75 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (четвъртите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 5=9/10=0,9 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (четвъртите три), $\Delta max.=1$, а $\lambda=1^2/2 \times 4=1/8=0,125 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (четвъртите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 6=9/12=0,75 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (петите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 5=9/10=0,9 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (петите три), $\Delta max.=2$, а $\lambda=2^2/2 \times 4=4/8=0,5 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (петите три), $\Delta max.=2$, а $\lambda=2^2/2 \times 6=4/12=0,33 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо шест килолекси текст, по три от *RZ* и *LSK* (шестите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 5=9/10=0,9 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (шестите три), $\Delta max.=1$, а $\lambda=1^2/2 \times 4=1/8=0,125 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо шест килолекси текст, по три от *RZ* и *LSK* (шестите три), $\Delta max.=3$, а $\lambda=3^2/2 \times 6=9/12=0,75 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

2.2.3.4.2. Критерий на Колмогоров-Смирнов за *LSK* и *CPS*

Във вариационна редица от номинални температури в общо десет килолекси текст - петте от *CPS* и първите пет от *LSK* $\Delta max.=1$, а $\lambda=1^2/2 \times 6=1/12=0,083 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо десет килолекси текст - петте от *CPS* и първите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 8=4/16=0,25 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо десет килолекси текст - петте от *CPS* и първите пет от *LSK* $\Delta max.=3$, а $\lambda=3^2/2 \times 10=9/20=0,45 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо десет килолекси текст - петте от *CPS* и вторите пет от *LSK* $\Delta max.=3$, а $\lambda=3^2/2 \times 8=9/16=0,56 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо десет килолекси текст - петте от *CPS* и вторите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 7=4/14=0,29 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо десет килолекси текст - петте от *CPS* и вторите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 8=4/16=0,25 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо десет килолекси текст - петте от *CPS* и третите пет от *LSK* $\Delta max.=4$, а $\lambda=4^2/2 \times 6=16/12=1,33 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо десет килолекси текст - петте от *CPS* и третите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 7=4/14=0,29 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо десет килолекси текст - петте от *CPS* и третите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 8=4/16=0,25 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо десет килолекси текст - петте от *CPS* и четвъртите пет от *LSK* $\Delta max.=3$, а $\lambda=3^2/2 \times 6=9/12=0,75 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо десет килолекси текст - петте от *CPS* и четвъртите пет от *LSK* $\Delta max.=2$, а $\lambda=2^2/2 \times 8=4/16=0,25 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо десет килолекси текст - петте от *CPS* и четвъртите пет от *LSK* $\Delta max.=4$, а $\lambda=4^2/2 \times 8=16/16=1 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

2.2.3.4.3. Критерий на Колмогоров-Смирнов за *LSK* и $\check{S}M$

Във вариационна редица от номинални температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и *LSK* (първите шест), $\Delta max.=4$, а $\lambda=4^2/2 \times 8=16/16=1 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и *LSK* (първите шест), $\Delta max.=6$, а $\lambda=6^2/2 \times 11=36/22=1,64 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и *LSK* (първите шест), $\Delta max.=5$, а $\lambda=5^2/2 \times 9=25/18=1,56 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и *LSK* (вторите шест), $\Delta max.=5$, а $\lambda=5^2/2 \times 11=25/22=1,14 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адекватни температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и *LSK* (вторите шест), $\Delta max.=6$, а $\lambda=6^2/2 \times 11=36/22=1,64 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от глаголни температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и LSK (вторите шест), $\Delta_{max}=5$, а $\lambda=5^2/2 \times 7=25/14=1,79 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от номинални температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и LSK (третите шест), $\Delta_{max}=5$, а $\lambda=5^2/2 \times 10=25/20=1,25 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Във вариационна редица от адеквативни температури в общо дванадесет килолекси текст, по шест $\check{S}M$ и LSK (третите шест), $\Delta_{max}=6$, а $\lambda=6^2/2 \times 8=36/16=2,25 > 1,84$, но $2,25 < 2,65$. Следователно би трябвало резултатите от съпоставката на двата вариационни реда да попадат „в границите на съмнението“.

Във вариационна редица от глаголни температури в общо дванадесет килолекси текст, по шест от $\check{S}M$ и LSK (третите шест), $\Delta_{max}=5$, а $\lambda=5^2/2 \times 8=25/16=1,56 < 1,84 < 2,65$. Следователно би трябвало двата вариационни реда да са от една генерална съвкупност.

Според направените изчисления по Критерия на Колмогоров-Смирнов показателите на LSK и при съществителните имена, и при прилагателните имена (резултатите от третите шест килолекси все пак попадат „в границите на съмнението“), и при глаголните форми би трябвало да принадлежат към една генерална съвкупност с показателите на RZ , $\check{S}M$ и CPS .

2.2.3.5. Критерий на Фишер⁷

Когато съпоставяме ексцерпт, състоящ се от проби, които имат статистически инвариант, с друг ексцерпт, състоящ се от проби, които също имат статистически инвариант, но друг, е възможно критерият λ да не регистрира различие, защото е възможно двата ексцерпта да се различават не по средните си аритметични, а по показателите си за разнообразието на вариантите, от които се състоят вариационните редове, т.е. по сигмите си, или по средните квадратични. За да се регистрира това различие съществува критерий, който се нарича Критерий на Фишер. Така ако сме изчислили средните квадратични отклонения на два вариационни реда от една езикова категория и разделим по-голямото на по-малкото, по големината на получената стойност можем да съдим дали двата вариационни реда действително се различават статистически. Ако полученото число е по-малко от 1,69, трябва да се признае наличието на един статистически инвариант за пробите от двата ексцерпта. Ако е по-голямо от 2,11, трябва да се приеме, че разликата между дисперсиите е достатъчно ярка и можем да разглеждаме ексцерптите като принадлежащи на различни от глотометрическо гледище генерални съвкупности (на различни по жанр текстове).

Тук ще направя съпоставка на резултатите от изчислението на вече познатата категория за трите килолекси RZ и шест групи от по три килолекси LSK , за петте килолекси CPS и четири групи от по пет килолекси LSK и за шестте килолекси $\check{S}M$ и три групи от по шест килолекси LSK .

2.2.3.5.1. Критерий на Фишер за LSK и RZ

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в RZ е 0,58; 1,53; 4,04. Средното квадратично

⁷ По-подробно за Критерия на Фишер, вж. Янакиев 1977: 210-212.

отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в първите три килолекси *LSK* е 1,73; 2,52; 7,77.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $1,73/0,58=2,98>2,11>1,69$, т.е. двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,52/1,53=1,65<1,69<2,11$, т.е. двата съпоставяни ексцерпта имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $7,77/4,04=1,92>1,69$, но $1,92<2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата ексцерпта попадат „в границите на съмнението“.

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми във вторите три килолекси *LSK* е 5,97; 2,65; 3,51.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $5,97/0,58=10,29>2,11>1,69$, т.е. двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,65/1,53=1,73>1,69$, но $1,73<2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата ексцерпта попадат „в границите на съмнението“.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,04/3,51=1,15<1,69<2,11$, т.е. двата съпоставяни ексцерпта имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в третите три килолекси *LSK* е 3,51; 1; 3,21.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $3,51/0,58=6,05>2,11>1,69$, т.е. двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $1,53/1=1,53<1,69<2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата ексцерпта имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,04/3,21=1,26<1,69<2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в четвъртите три килолекси *LSK* е 3,79; 2,52; 3,22.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $3,79/0,58=6,53>2,11>1,69$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,52/1,53=1,65 < 1,69 < 2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата ексцерпта имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,04/3,22=1,25 < 1,69 < 2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в петите три килолекси *LSK* е 6,03; 0,58; 4,51.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $6,03/0,58=10,4 > 2,11 > 1,69$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $1,53/0,58=2,64 > 2,11 > 1,69$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,51/4,04=1,12 < 1,69 < 2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в шестите три килолекси *LSK* е 2,08; 1,53; 1,53.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $2,08/0,58=3,59 > 2,11 > 1,69$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $1,53/1,53=1 < 1,69 < 2,11$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,04/1,53=2,64 > 2,11 > 1,69$, т.е. по отношение на съотношението между средните си квадратични отклонения двата съпоставяни ексцерпта нямат общ статистически инвариант.

Въз основа на направените изчисления можем да обобщим, че: по отношение на вариационните редове от съществителни имена двата съпоставяни текста нямат общ статистически инвариант между нито една шестте проби от *LSK* и пробата от *RZ*. По отношение на вариационните редове от прилагателни имена двата съпоставяни текста по-скоро имат общ статистически инвариант - четири пъти имат, веднъж попадат в „границите на съмнението“, веднъж нямат. По отношение на вариационните редове от глаголни форми двата съпоставяни текста по-скоро имат общ статистически инвариант - четири пъти имат, веднъж попадат в „границите на съмнението“, веднъж нямат.

2.2.3.5.2. Критерий на Фишер за LSK и CPS

Средното квадратично отклонение за вариационните редове глаголни форми, съществителни имена и прилагателни имена в петте килолекси CPS е: 3,21; 2,61; 2,39. Средното квадратично отклонение за тези категории в първите пет килолекси LSK е: 6,04; 3,78; 3,29.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $6,04/3,21=1,88 > 1,69$, но $< 2,11$, т.е. в „границите на съмнението”.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $3,78/2,61=1,45 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $3,27/2,39=1,37 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове глаголни форми, съществителни имена и прилагателни имена във вторите пет килолекси LSK е: 4,44; 6,22; 2,3.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,44/3,21=1,38 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $6,22/2,61=2,38 > 2,11 > 1,69$, т.е. двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,39/2,3=1,04 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове глаголни форми, съществителни имена и прилагателни имена в третите пет килолекси LSK е: 4,04; 5,17; 1,87.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $4,04/3,21=1,26 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $5,17/2,61=1,98 > 1,69$, но $1,98 < 2,11$, т.е. „в границите на съмнението”.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,39/1,87=1,28 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Средното квадратично отклонение за вариационните редове глаголни форми, съществителни имена и прилагателни имена в четвъртите пет килолекси LSK е: 1,14; 1,73; 3,11.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $3,21/1,14=2,82 > 2,11 > 1,69$, т.е. двата съпоставяни ексцерпта нямат общ статистически инвариант.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $2,61/1,73=1,51 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $3,11/2,39=1,3 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Въз основа на направените изчисления можем да заключим, че: по отношение на вариационните редове от съществителни имена е възможно двата съпоставяни текста по-скоро да имат общ статистически инвариант (два пъти имат, веднъж попадат в „границите на съмнението“, веднъж нямат), по отношение на вариационните редове от прилагателни имена двата съпоставяни текста по-скоро имат общ статистически инвариант (четири пъти имат), по отношение на вариационните редове от глаголни форми двата съпоставяни текста по-скоро имат общ статистически инвариант (два пъти имат, веднъж попадат в „границите на съмнението“, веднъж нямат).

2.2.3.5.3. Критерий на Фишер за LSK и ŠM

Средното квадратично отклонение за вариационните редове съществителни имена, прилагателни имена и глаголни форми в шестте килолекси ŠM е: 6,92; 2,25; 5,53. Средното квадратично отклонение за тези категории в първите шест килолекси LSK е: 3,69; 3,45; 4,54.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $6,92/3,69=1,88 > 1,69$, но $1,88 < 2,11$, т.е. в „границите на съмнението“.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $3,45/2,25=1,53 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $5,54/5,53=1 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Средното квадратично отклонение за тези категории във вторите шест килолекси LSK е: 4,26; 3,46; 3,15.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $6,92/4,26=1,62 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $3,46/2,25=1,54 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $5,53/3,15=1,76 > 1,69$, но $1,76 < 2,11$ т.е. „в границите на съмнението“.

Средното квадратично отклонение за тези категории в третите шест килолекси LSK е: 4,98; 1,17; 3,06.

Стойността на Критерия на Фишер за съществителните имена в съпоставяните ексцерпти е $6,92/4,98=1,39 < 1,69 < 2,11$, т.е. би трябвало двата съпоставяни ексцерпта да имат общ статистически инвариант.

Стойността на Критерия на Фишер за прилагателните имена в съпоставяните ексцерпти е $2,25/1,17=1,92 > 1,69$, но $1,92 < 2,11$, т.е. „в границите на съмнението“.

Стойността на Критерия на Фишер за глаголните форми в съпоставяните ексцерпти е $5,53/3,06=1,81 > 1,69$, но $1,81 < 2,11$, т.е. „в границите на съмнението“.

Въз основа на направените изчисления можем да заключим, че: по отношение на Критерия на Фишер за вариационните редове от съществителни имена е възможно двата съпоставяни текста по-скоро „да имат общ статистически инвариант“ - два пъти имат, веднъж попадат „в границите на съмнението“. За вариационните редове от прилагателни имена двата съпоставяни текста по-скоро „имат общ статистически инвариант“ - два пъти имат, веднъж попадат в „границите на съмнението“. За вариационните редове от глаголни форми двата съпоставяни текста по-скоро „имат общ статистически инвариант“ - веднъж имат, два пъти попадат в „границите на съмнението“.

Въз основа на направените изчисления по Критерия на Фишер можем да обобщим, че *LSK* количествено по-скоро има общ статистически инвариант *cRZ, CPS*, и *ŠM* по посочените показатели. Прави впечатление, че колкото по-големи по обем стават книжовните паметниците, с които се съпоставя, толкова по-разнообразни стават и резултатите, а това означава, че те важат преди всичко за съпоставяните ексцерпти, а чак след това и за текстовете, на които те принадлежат и е възможно да се влияят както от обема на съпоставяните паметници, така също и от жанра (особено след като някои от разгледаните текстове проявяват особености, характерни за повече от един средновековен жанр).

3. Заключение

В настоящото изследване бяха представени данните от глотометричния анализ на едно от най-ранните известни засега средновековни хърватски произведения, писано на народен език и записано с латински букви, житието на св. Екатерина *Legenda sv. Katarine divice* от ръкопис *Ic6* от Архива на ХАЗУ. Направени бяха следните обобщения и заключения:

3.1. Показатели на словника

Въз основа на изнесените резултати, въпреки малкия обем на паметника, може да се забележи тенденцията да е висок процентът на служебната лексика сред най-често употребяваните думи в текста. Колкото повече намалява честотата на употребените единици в текста, намалява и количеството на служебната лексика, а нараства това на пълнозначната. Относително niskият процент на служебната лексика в словника показва, че голяма част от единиците тук са употребени по повече от веднъж в текста.

Отношението *покритие в текста - покритие в словника* показва, че колкото по-висока е честотата на употреба на лексемите, по-голямо е и покритието в текста, а по-малко - в словника. С намаляването на честотата на употреба, намалява покритието на групата в текста, а нараства покритието в словника.

3.2. Показатели на отношението между словника и текста

3.2.1. Коефициент на лексикално многообразие

Приведените резултати все пак показват известна стабилност за разглежданата категория в най-ранните хърватски текстове, писани на народен език и записани с латински букви, и са свидетелство за това, че макар и да не се опират на предходна

книжовна традиция от подобен характер, хърватските книжовници все пак успяват да създадат текстове, които дори и преводни, да отразяват особеностите и на говоримия език, но и да са лексикално разнообразни. Все пак това са различни по обем текстове, така че приведените данни биха могли да служат само като ориентир и отправна точка при бъдещи съпоставки на равни части от тези и други подобни произведения.

3.2.2. Индекс на повторемостта

Сведенията за резултатите от други средновековни хърватски текстове, писани на народен език с латински букви, показват относително ниски стойности на разглежданата категория в ранната хърватска литература, което показва, че вероятно и тя е относително стабилна в текстовете на народен език от този период. Все пак при някои от произведенията ниските стойности може да се дължат и на по-малкия обем на паметника.

И тук трябва да уточним, че това са различни по обем текстове, така че приведените данни биха могли да служат само като ориентир и отправна точка при бъдещи съпоставки на равни части от тези и други подобни произведения. Напълно очаквано е колкото по-ниска е стойността на индекса на повторемост, толкова по-голямо да е разнообразието на речника на дадено произведение.

3.3. Показатели на текста

3.3.1. Статистически изчисления

Това са различни по обем текстове, така че приведените данни биха могли да служат само като ориентир и отправна точка при бъдещи съпоставки на равни части от тези и други подобни произведения. Все пак напълно е възможно някои от изчислените категории да се влияят и от жанровите специфики на произведенията.

3.3.2. Съпоставка на ексцерпти от други текстове

Получените резултати не дават еднозначен отговор. Тъй като в лингвостатистиката се съпоставят на първо място ексцерпти, а чак след това чрез тях и текстовете, на които те принадлежат, трябва да уточним, че направените обобщения в съответните раздели все пак важат преди всичко за ексцерптите.

Според направените изчисления по Критерия на Колмогоров-Смирнов показателите на *LSK* за всички свои части би трябвало да принадлежат към една генерална съвкупност показателите на *RZ*, *ŠM* и *CPS*.

Въз основа на направените изчисления по Критерия на Фишер можем да обобщим, че *LSK* количествено по-скоро има общ статистически инвариант с *RZ*, *CPS*, и *ŠM* по посочените показатели. Прави впечатление, че колкото по-големи по-обем стават паметниците, с които се съпоставя, толкова по-разнообразни стават и резултатите, а това означава, че те важат преди всичко за съпоставяните ексцерпти, а чак след това и за текстовете, на които те принадлежат и е възможно да се влияят както от обема на съпоставяните паметници, така също и от жанра (особено след като някои от разглежданите текстове проявяват особености, характерни за повече от един средновековен жанр).

ЛИТЕРАТУРА

- Давидов 1996: *Давидов, А.* Използване на статистическите методи в лексиколожките изследвания. – В: Старобългарска лексикология. Под. ред. на А. Давидов. Велико Търново, 28-43.
- Илиева 2011: *Илиева, Т.* Лексиката в Йоан-Екзарховия превод на Богословието – количествени показатели. – В: Старобългаристика, XXXV, 2, 14-37.
- Илиева 2013: *Илиева, Т.* Лексиката в творчеството на Петър Черноризец – глотометричен профил. – В: Език и литература, 1-2, 72-86.
- Стефанов 2014: *Стефанов С. С.* О стихе в средновековном житии св. Екатерины на народном хорватском языке. – В: Сборник тезисов двух Международных молодёжных междисциплинарных славистических конференций. Роль славянской молодёжи в процессе устойчивого цивилизационного развития. V. Славянский мир: аспекты истории и культуры (28.05.2014), Москва, 40-44.
- Стефанов 2015а: *Стефанов, С.* Стилистични похвати за създаване на ритъм в средновековно житие на св. Екатерина на народен хърватски език. – В: Сборник с доклади от Дванадесетите международни славистични четения „Движение и пространство в славянските езици, литератури и култури” (9-10.05.2014 г.), София, 46-53.
- Стефанов 2015б: *Стефанов, С.* Лексика и стил в средновековно житие на св. Екатерина на народен хърватски език. – *Littera et Lingua*, Пролет-Лято. – <http://slav.uni-sofia.bg/naum/lilijournal/2015/12/1-2/>
- Стефанов 2018а: *Стефанов, С.* Лексиката в *Red i zakon zadarskih dominikanki*: Глотометричен анализ. – В: Сборник от XVI Международни Славистични четения „Стереотипът в славянските езици, литератури и култури“ 26-28.04.2018 г., (под печат).
- Стефанов 2018б: *Стефанов, С.* Лексиката в *Šibenska molitva*: Глотометричен анализ. Език и литература, (под печат).
- Стефанов 2018в: Лексиката в *Cantilena pro sabatho*: Глотометричен анализ. Филологически форум (под печат).
- Фрумкина 1964: *Фрумкина, Р. М.* Статистические методы изучения лексики. Москва.
- Христова 1994: *Христова, И.* Речник на словата на Климент Охридски. София.
- Янакиев 1977: *Янакиев, М.* Стилистиката и езиковото обучение. София: Държавно издателство „Народна просвета”,
- Badurina Stipčević 2013: *Hrvatska srednjovjekovna proza I. Legende i romani. Priredila i transkribirala Vesna Badurina Stipčević.* Zagreb: Matica Hrvatska, 106–114.
- Fališevac 1978: *Fališevac, D.* Razdoblja i književne vrste hrvatske srednjovjekovne proze (o problemu periodizacije i klasifikacije). – В: *Filologija*, 8, 89-98.
- Fališevac 2010: *Fališevac D.* Sveta Katarina – prva učena žena u hrvatskoj književnoj kulturi. – В: *Slovo* (60), 257-277.
- Ganza-Aras 1990: *Ganza-Aras, T.* Zadužbina i licej Sv. Dimitrija u Zadru. – В: *Historijski zbornik*, god. XLIII (1), 197-208.
- Hadrovics 1984: *Hadrovics, L.* Cantilena pro sabatho. – В: *Filologija*, 12. Zagreb, 7-25.
- Hamn 1926/1927: *Hamn, J.* Legenda o sv. Katarini iz Aleksandrije u našim starim rukopisima (дипломна работа, в ръкопис), (non vidi).
- Hercigonja 1973: *E. Hercigonja*, Hrvatska književnost srednjega vijeka (do kraja 14. stoljeća). – В: *Zbornik Zagrebačke slavističke škole 1.* Zagreb, 22-66.

- Hercigonja 1975: *Hercigonja, E.* Srednjovjekovna književnost. – B: Povijest hrvatske književnosti, knj. 2, Liber-Mladost. Zagreb, 38-39.
- Jagić 1869: *Jagić, V.* Ogledi stare hrvatske proze. II. Legenda sv. Katarine divice. – B: Starine, knj. 1, JAZU. Zagreb, 224-227.
- Malić 2008: *Malić D.* Croatian Medieval texts in Latin script. – B: Croatia and Europe, vol 2, Croatia in the Middle Ages and the Renaissance (A Cultural Survey). Zagreb, 301-322.
- Štefanić 1969: *Štefanić, Vj.i suradnici.* Hrvatska književnost srednjega vijeka. Pet stoljeća hrvatske književnosti (1). Zagreb, 274-277.