



UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

***I reč Cěsar'***

Titlo-abbreviations in Old East Slavic

Jens Kristian Skjølsvold

Master's thesis in Russian, RUS-3920, May 2023



# Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>1</b>
1.1	Reservations.....	1
1.2	The subject matter of this thesis. ....	1
1.3	Simplified research question .....	2
1.4	Means to an end.....	2
1.5	Disposition.....	3
<b>2</b>	<b>(Old) Church Slavonic and titlos .....</b>	<b>4</b>
2.1	Titlos.....	4
2.1.1	Idols and inverse meaning.....	6
<b>3</b>	<b>Development of Slavic Language .....</b>	<b>7</b>
3.1	Glagolitic & Cyrillic.....	7
3.2	Old East Slavic .....	9
3.2.1	Handwriting.....	9
3.2.2	Writing materials.....	14
3.3	Standardization of Russian .....	15
<b>4</b>	<b>Theory .....</b>	<b>16</b>
4.1	Tsar & God .....	16
4.1.1	Semiology.....	17
4.2	Diglossia.....	17
4.3	Cognitive Linguistics.....	20
4.3.1	The Cornerstones of Cognitive Linguistics.....	21
4.3.2	Metonymy and metaphor .....	24
4.3.3	Defining Metonymy .....	24
4.3.4	Polysemy .....	26
<b>5</b>	<b>Research Question(s).....</b>	<b>27</b>
<b>6</b>	<b>Data.....</b>	<b>28</b>

6.1	Digitalization and electronic corpora .....	28
6.2	The Tromsø Old Russian and OCS Treebank .....	28
6.2.1	The data in this thesis. ....	29
<b>7</b>	<b>Methodology .....</b>	<b>31</b>
7.1	Best Practice .....	31
7.2	Statistical models and tests .....	31
7.2.1	Correlation & Regression.....	31
7.2.2	Alternatives to Regression .....	33
7.3	Issues with Old Russian and R Markdown.....	33
7.4	R Markdown packages .....	33
<b>8</b>	<b>Statistical Analysis.....</b>	<b>34</b>
8.1	Levels.....	34
8.2	CART-Analysis & Random Forest.....	36
8.2.1	Prediction accuracy CART-Analysis & Random Forest.....	40
8.3	Generalized Logarithmic Regression Model(s), ANOVA & AIC .....	42
8.3.1	Prediction accuracy Generalized Logarithmic Regression Model.....	47
8.4	Outliers .....	47
8.5	Interpretation of the statistical analysis .....	49
<b>9</b>	<b>Discussion.....</b>	<b>49</b>
9.1	Meaning .....	50
9.2	Radial Network.....	56
9.3	Further research .....	59
<b>10</b>	<b>Conclusion.....</b>	<b>59</b>
	<b>Works cited .....</b>	<b>61</b>

## List of Tables

Table 1 - Nouns .....	29
-----------------------	----

## List of Figures

Figure 1 – Tsar abbreviated with a titlo. ....	2
Figure 2 - Gammanovich's list of common titlo-abbreviations.....	5
Figure 3 - Glagolitic .....	8
Figure 4 - (Old) Church Slavonic Alphabet .....	9
Figure 5 - Ustav.....	10
Figure 6 - Poluustav .....	11
Figure 7 - Skoropis' .....	13
Figure 8 – The Network Commitment .....	23
Figure 9 - Typology of metonymical relations.....	26
Figure 10 - CART analysis.....	38
Figure 11 - Random Forest.....	40
Figure 12 - Influence Plot.....	48
Figure 13 - Radial Network.....	58

## Foreword

First and foremost, I would like to thank my supervisor Laura Alexis Janda, and Hanne Eckhoff who compiled the datasets used in this thesis. I would also like to extend my thanks to Tore Nettet and Svetlana Sokolova for facilitating courses so that I have been able to pursue my personal interests in all that is *Rus'*, ranging from *skomorokhs* and Muscovite court customs (such as *bit' chelom*) to diacritic marks throughout my studies at the master's degree program in Russian.

# 1 Introduction

## 1.1 Reservations

In this thesis I oftentimes speak of Old East Slavic when I refer to the ancestor language(s) of contemporary modern Russian. Old East Slavic should herein be considered an umbrella term that spans both Old Russian and Middle Russian. Moreover, I use the terms Rus', Medieval Russia, and Muscovy somewhat interchangeably to address historical Russia in terms of both the geographical area and the state.

Readers of this thesis are presumably not only familiar with contemporary modern Russian, but also Old Church Slavonic and Old East Slavic, which allows me to take certain liberties. Old Church Slavonic and Old East Slavic words are written in Cyrillic and translated, though without transliterations, while modern Russian is largely transliterated, however not translated.

The thesis includes a statistical analysis first submitted for examination in the course HIF-3082: Quantitative Methods in Linguistics. Most of the text from the original examination paper has been reworked to fit this thesis, though the statistical analysis and associated prose presented in section 8 of this thesis is largely untouched, except for minor corrections.

## 1.2 The subject matter of this thesis.

In my bachelor's thesis in *Russlandsstudier* (Russian studies) I wrote about the religious sanction behind tsar Ivan IV the Terrible's reign, and particularly the state policy and time known as *Oprichnina* (1565-1572), which was characterized by mass-persecutions, violence and repression of the general population, aristocracy, and clergy of Muscovy. *Oprichnina* would eventually internally displace most of the *Rusian* population. What was curious about Ivan's reign, was that his autocracy was not institutionalized by brute force alone, but also through religious sanction and juxtaposition of the muscovite tsar with God through the tsar's power, which was ordained by God according to the Muscovites.

I read about the peculiarities of the role of the tsar (and about Ivan IV) in Victor Zhivov & Boris Uspenskij's *Tsar and God and Other Essays in Russian Cultural Semiotics* (2012). This work touched upon so-called titlos, an orthographical feature, which was used in conjunction with abbreviations to testify to the inherent divine attributes and nature of an object – in this case, the muscovite monarch. While Zhivov & Uspenskij did not provide an example, I have provided one in figure 1 below, where tsar is abbreviated by titlo:

Figure 1 – Tsar abbreviated with a titlo.



At that time, I was not concerned with linguistics, so I did not delve deeper into titlos. However, some months later, during readings in class throughout the course RUS-3010 *Eldre språk og litteratur*, several texts in both Old Church Slavonic and Old East Slavic contained selected abbreviated words, which had this strange dash above the letters, akin to a diacritic, but different. Upon translation, it became clear that many of these words were religious, or more precisely, sacred objects, whilst others were numbers or dates. To some extent, it seemed codified, however, the fuzziness of titlo-abbreviations motivated the research questions posed in this thesis.

### 1.3 Simplified research question

In the simplest form, this thesis asks *why* did Old East Slavic scribes use titlo-abbreviations?

### 1.4 Means to an end.

Luckily, it is not my task to make novel breakthroughs in neither theory nor methodology in this thesis. A fitting paradigmatic approach to these research questions is presented by Janda & Dickey in *Cognitive Linguistics* (2017), where they describe the applicability of cognitive linguistics theory and statistical methods for research in Slavic linguistics.

In general, cognitive linguistics has focused on analyzing the semantic categories that make up meaning in linguistic units. In terms scope of cognitive linguistics, units come in varying levels of complexity: from the word-internal units to discourse level units, described as constructions, and their pattern of use. Moreover, cognitive linguistics views language as an emergent structure, which puts more emphasis on similarities, than differences. Cognitive linguistics does not reduce language to minimal units of composition.

Cognitive linguistics flourished at the same time as research materials became readily available through electronic linguistic corpora, with the help of statistical analysis, and cognitive linguistics has since assumed a leading role in the interpretation of statistical outcomes in linguistic data. Oftentimes, statistics have testified to variation where linguistics have traditionally spoken of rules. In cognitive linguistics, the meaning of a linguistic expression or any meaningful structure revolves around a prototype and extended

relationships, often motivated by metaphor or metonymy. Radial categories have been extensively used to analyze Slavic lexical and grammatical categories. Research into the semantic meaning of grammatical and lexical categories in the scope of cognitive linguistics has provided new insights into major Slavic grammatical categories over the course of 30 years. Cognitive linguistics is adept at inquiry into lexico-grammatical units, while it also describes and explains variation in meaning across time and speakers through statistical methods.

Applying cognitive linguistics framework to Old East Slavic is not a novel approach, as it was applied, for example, by Eckhoff and Janda in 2014 to demonstrate that Old Church Slavonic verb forms could be sorted by aspect in early Slavic texts. Given the applicability of cognitive linguistics to Slavic in general, and its demonstrated use for inquiry into Old Church Slavonic, it should readily be applicable for researching titlos in Old East Slavic texts. Moreover, cognitive linguistics' compatibility with statistical methods dictates that quantitative inquiry into the concept may shed new light on the meaning of the titlos.

## 1.5 Disposition

To answer the research question posed above, I will start by diving into the nature of the concept in question, that is, titlos and titlo-abbreviations. However, it becomes apparent as this thesis unfolds that the available literature on titlos concerns itself with (Old) Church Slavonic, whereas this thesis concerns itself with titlos in Old East Slavic. Therefore, I will go through the development of Slavic Language and Slavic paleography, handwriting in *Rus'*, along with the standardization of Russian, to contextualize the concept in *Rus'*.

This brings us to the theory section of thesis, where I reiterate Zhivov & Uspenskij's theoretical notions about titlos, while briefly addressing semiology. I then explicate upon the theory of diglossia in *Rus'*, which revolves around the dynamics between (Old) Church Slavonic and Old East Slavic in *Rus'*. This is followed by Cognitive Linguistics theory. Cognitive Linguistics literature is a vast field; therefore, I will only go in depth on metonymy and polysemy, besides explicating upon the cornerstones of Cognitive Linguistics. With a firm understanding of the concept in question and its context, along with theory to frame it in, I will specify the research question(s) posed above in greater detail.

What comes next is two exhaustive chapters addressing the digitalization of linguistics research, corpora, and methodology - both in general and the specifics concerning this thesis.



The two subsequent chapters consists of a statistical analysis, and the interpretation of the results from the statistical analysis in the discussion. At the end of the discussion, I provide a suggestion for a radial network of related meanings, concerning titlo-abbreviations.

## 2 (Old) Church Slavonic and titlos

I must stress that the section that follows concerns itself with titlos in (Old) Church Slavonic, while the questions raised in this thesis, concerns itself with titlos in Old East Slavic.

### 2.1 Titlos

In (Old) Church Slavonic particular words may be abbreviated, or intentionally be written without certain letters, while a titlo or the omitted letter in miniature is written above the word, as described by Alypy Gammanovich in his authoritative work *Grammar of the Church Slavonic Language* (2001). This thesis concerns itself with the *simple titlo* [˘] (demonstrated in figure 1 above), however, I will address the *lettered titlos* that appear in miniature briefly. The letters that commonly appear in miniature are с, г, д, о, and р, that is, *slovo*-titlo, *glagol*'-titlo, *dobro*-titlo, *on*-titlo, and *rtsy*-titlo, respectively (see figure 2 below, where several of the listed words include *lettered titlos*). Titlos can also be utilized above letters to denote numbers in manner like that of Greek, however, numbers are not matter of concern in this paper. The *simple titlo* [˘], will henceforth be addressed as titlo, unless context suggests otherwise.

As specified, it was not just any word that could be abbreviated by titlo. Only objects that are particularly respected or revered can be abbreviated by titlos. Furthermore, the use of a titlo and the lack of one attest to semantic differences. Observe, for example, богъ (God) without a titlo and Б҃ъ with a titlo. The former speaks of an idol or pagan deity, while the latter refers to the One, the true God. In his work, Gammanovich provides a list of words usually abbreviated by a titlo (see figure 2 below). The list includes archetypal Christian Biblical figures or their metaphorical extension, such as царь (tsar) - црѣ, сынъ (son) - Снѣ, and отецъ (father) - оцѣ, but also words such as молитва (prayer) - млѣтва,

Figure 2 - Gammanovich's list of common titlo-abbreviations

А́ггъ [angel—angel]	ИѢ́РУСА́ЛМЪ [Jerusalem — Jerusalem]	МА́РІА [María — Mary]	Трѣ́ца [Troítsa — Trinity]
А́пстолъ [apóstol—apostle]	И́исусъ [Iisús — Jesus]	МА́ТН [Máti — Mother]	Хрѣ́стовъ [Hristós — Christ]
Бо́гъ [Bog — God]	Крѣ́тъ [krest — Cross]	МА́ЛТВА [molítva — prayer]	Цѣ́рковь [tsárstvo — kingdom]
Божѣ́ственный [bozbestvenny — divine]	Крѣ́стичель [Krestitel' — the Baptist]	МА́ЛОСТЬ [mílost' — mercy]	Царь [tsar' — king]
Бла́гъ [blag — good]		МА́ЛОСТІЕ [milosérdie — kindheartedness]	Црѣ́ковъ [tsérkov' — church]
Бла́женъ [blazbén—blessed, blissful]		МА́ЛЕНЦЪ [mladénets—Infant]	Чѣ́стнѣ́й [chestnýj — honourable]
Благосло́венъ [blagoslovén — blessed]		МА́ЧНИКЪ [múchenik — martyr]	Чѣ́стнѣ́й [čbístyj] [clean, pure]
Благочѣ́стно [blagochéstno — devoutly]		НЕ́бо [nébo — heaven]	
Бла́гость [blagodát' — grace]		Нѣ́дѣ́ля [nedélja—Sunday]	
Бѣ́га [Bogoróditsa— Theotokos]		О́тъцъ [Otéts — Father]	
Воскрѣ́ніе [voskresénie—resurrection]		Правѣ́дникъ [právednik — righteous man]	
Вла́да [Vladýka — Master]		Правѣ́денъ [pre podóben — venerable]	
Вла́дица [Vladýchitsa — Lady]		Прѣ́сто́лъ [prestól — Throne; Holy Table]	
Гос́дь [Gos pód' — Lord]		Проро́къ [prorók — prophet]	
Дѣ́ва [Děva — Virgin]		Свѣ́тъ [svjat — holy]	
Духъ [Dukh — Spirit]		Свѣ́титель [svjattitel' — hierarch]	
Еписко́пъ [epískop — bishop]		Спасъ [S pas — Saviour]	
Евангѣ́ліе [Evanǵelie — Gospel]		Сынъ [syn — Son]	
И́мярекъ [imjarek — “supply proper name”]			

(Gammanovich, 2001, p. 29-30)

In ПРАВОСЛАВНЫЙ БОГОСЛУЖЕБНЫЙ ТЕКСТ И СОЦИАЛЬНЫЕ АСПЕКТЫ ФУНКЦИОНИРОВАНИЯ ПРАВОСЛАВНОГО ПРЕЛИГИОЗНОГО СОЗНАНИЯ (2006)

Sazonova details the semantics of titlo-abbreviation, based upon Gammanovich's grammar (amongst others). As with the differences between Бѣ́гъ and богъ explained above, ангелъ written without a titlo can refer to an evil spirit, whereas а́ггъль speaks of an angel of God. While dependent on overall context, we may state, in short, that the titlo refers to that which is sacred, blessed, or holy.

For instance, человек (man) would often be written under a titlo as члѣ́къ in older works. While written out in its whole, it often points to the sinful man who seeks penitence through prayer and confession. Furthermore, when человек is written under a titlo, it also alludes to mankind as being created in God's image, and furthermore to God's love for his creation, as it is seen through the eyes of the Lord. We may then assume, that человек written out letter by letter, alludes to man in the scope of original sin.

In essence, a given word written under a titlo points to the sacral features of the object, while without a titlo it points to the semantic, or semiotic, inverse. As such, one must be familiar with both the writings, prayers, and the faith itself to command complete understanding of the

meaning of titlos. This in turn should emphasize the meaning, but also the reality of what is written, and the colossal boundaries in between mankind and the divinity of God, which separates them - the believers and God.

As briefly mentioned in terms of молитва, titlos are not reserved for nouns (or proper names) and their metaphorical extensions (as in богъ, хрїстось [Christ], богородица [Mother of God], and воскресенїе [Resurrection]). They are also used above words such as милость (mercy), милостивый (merciful), and чистый (pure), which gives special meaning to these words in the eyes of the reader.

### **2.1.1 Idols and inverse meaning**

According to the literature above, titlos index a special meaning, but we're also presented with the curious case, that a lack of titlo supposedly may invoke the inverse meaning, but it's not necessarily as simple as looking up the antonyms of words like богъ, сынъ or человек.

From a strictly Judeo-Christian perspective one might assume that the inverse of meaning of God and Christ would be biblical idols such as Baal or the Devil, and the Antichrist, respectively. However, the religious landscape of medieval *Rus'* was, arguably, a much more complex one.

Ivanov (*Dvoeverie*, 2010) writes, that despite Vladimir the Great's later adoption of Eastern Orthodoxy, many inhabitants of *Rus'* continued to worship the old Slavic pantheon, a phenomenon known as *dvoeverie*, or dual faith. As such, the deities of the old pantheon were given the name of Christian saints, which led to a syncretization of the old and new faith (c.f. Elijah depicted whilst riding a chariot in the sky like *Perun*, St. George fighting a dragon reminiscent of *Perun* fighting the serpent. Moreover, depictions of St. Basil show domestic animals looking to him, in juxtaposition to the pagan god *Veles'* role in the old native belief. Ivanov links this syncretization of the pagan pantheon and Christianity with the purported diglossia of *Rus'* (which is addressed in section 4.3).

Furthermore, Kozlov and Matveeva (2021) maintain that we cannot necessarily trust (ancient) written sources as they most certainly were biased to the contemporary worldly or political reality of their time, and that even chronicle authors displayed tendentious political views. Archeological evidence, though, can complement (or distort) written sources. Nestor, the chronicler accredited as the author of *Povest vremennikh let'*, named several pagan deities worshipped in the pantheon established by Vladimir the Great (of Kiev) at his ascension to

power (980 A.D.), such as Перун, Хорс, Дажьбог, Стрибог, Макош, and Семаргл, whereas other Old Russian chronicles also mention Сварог and Велес, all of whom are attested by archeological evidence – oftentimes in the close proximity of churches. Curiously, archeological findings also include unidentified two-faced idols.

Given this situation, it is not a straightforward process of deciphering what idols or which pagan deity scribes and authors referred to. It is not within the scope of this thesis to address the topic of *dvoeverie*, false gods and such further, but it does highlight some of the problems, or nuances, historical or diachronic linguistics are faced with.

### **3 Development of Slavic Language**

Having addressed titles themselves and detailed their meaning in (Old) Church Slavonic, it is time to address the development of Slavic Language in general. Tore Nessel gives a short and concise timeline of the development of Slavic in language in *How Russian came to be the way it is* (2015, p. 10-11), which is described below.

The development from Proto-Slavic language (a progeny of Proto-Indo-European) to contemporary Modern Russian can be divided into the five periods: First, the period of Pre-Slavic (until 300 AD) before Slavic differed from other Indo-European languages. Second, the period of Common Slavic (300-1000 AD), when all the Slavic languages went through the same changes. Third, the period of Old Russian (1000-1400 AD), the namesake, which was the common ancestor language of modern Russian, Belarusian and Ukrainian. Fourth, the period of Middle Russian (1400-1700 AD), which was when Russian separated from what became Belarusian, and Ukrainian, and fifth, the period of Modern Russian (1700-present).

While categorizing the development of Slavic language in such brackets we must keep in mind that changes and development in language did not occur abruptly, but fluently. As such, the time frames attributed to the various iterations of what came to be Russian are only approximations.

#### **3.1 Glagolitic & Cyrillic**

As described by Nessel (*How Russian came to be the way it is*, 2015), Cyril and Methodius devised an alphabet to teach the New Testament in Slavic, and thus, as far as we know, introduced the art of writing to the Slavic World. This new alphabet was neither Latin nor Cyrillic, but Glagolitic. Whereas the Cyrillic alphabet was adopted from the Greek alphabet,

the origins of Glagolitic are hypothesized as either 1) being created from scratch, 2) derived from Greek cursive, or 3) based on Cyrillic. However, the 3<sup>rd</sup> hypothesis can be discarded entirely according to Nessel, as evidence simply suggests that Glagolitic is older than Cyrillic, while the 1<sup>st</sup> and 2<sup>nd</sup> hypotheses largely overlap.

Figure 3 - Glagolitic



(Gammanovich, 2001, p. 10)

The canon of Cyril and Methodius (along with their followers) is known as Old Church Slavonic (OCS), or старославянский язык in Russian. OCS encompasses texts from the 10<sup>th</sup> century, whereas Church Slavonic refers to texts dated after 1100 with several dialectal features, all however written in the Cyrillo-Methodian tradition (Nessel, 2015). See figure 4 below for an overview of the letters of (Old) Church Slavonic, provided by Gammanovich (2001, p. 22):

Figure 4 - (Old) Church Slavonic Alphabet

Ѧ, Ѧ—az [a]	Ѧ, Ѧ—tverdo [t]	Ѧ, Ѧ—yus bol'shóy
Ѣ, Ѣ—búki [b]	Ѧ, Ѣ, ѣ—uk [u]	ѦѢ, ѢѢ—yus bol'shóy iotírovanny
Ѥ, Ѥ—védi [v]	Ѧ, Ѥ—fert [f]	ѦѢ, ѢѢ—yus mály iotírovanny
Ѧ, Ѧ—glagól' [g, gb]	Ѧ, Ѧ—kher [kb—cf. 'loch']	
Ѧ, Ѧ—dobró [d]	Ѧ, Ѧ—ot [o + t]	
Ѥ, Ѥ, Ѥ—yest' [e, ye]	Ѧ, Ѥ—tsy [ts]	
Ѧ, Ѧ—zhivéte [zb, = s in pleasure]	Ѧ, Ѥ—cherv' [ch in church]	
Ѥ, Ѥ—zeló [z]	Ѧ, Ѥ—sha [sh]	
Ѥ, Ѥ—zeml'á [z]	Ѧ, Ѥ—shcha [shch]	
Ѧ, Ѧ—ízhe [i in 'machine']	Ѧ, Ѧ—yer [hard mark; silent]	
Ѧ, Ѧ—i [i]	Ѧ, Ѥ—yerý [similar to i in 'bit']	
Ѧ, Ѧ—káko [k]	Ѧ, Ѥ—yer' [soft mark; silent]	
Ѧ, Ѧ—liúdi [l]	Ѧ, Ѥ—yat' [ye]	
Ѧ, Ѧ—mysléte [m]	Ѧ, Ѥ—yu [yu; Eng. u in use]	
Ѧ, Ѧ—nash [n]	Ѧ, Ѥ—ya [ya]	
Ѧ, Ѧ, Ѧ—on [Eng. 'more']	Ѧ, Ѥ—mály yus [ya]	
Ѧ, Ѧ—omega [o]	Ѧ, Ѧ—ksi [x in express]	
Ѧ, Ѧ—pokóy [p]	Ѧ, Ѧ—psi [ps]	
Ѧ, Ѧ—rtsý [r trilled]	Ѧ, Ѧ—fitá (i.e. Greek theta) [f]	
Ѧ, Ѧ—slóvo [s]	Ѧ, Ѧ—ízhitsa [i]	

## 3.2 Old East Slavic

In *The Russian Language before 1700*, Matthews (1953) states that Old East Slavic traces its origins to Old Church Slavonic in 1056 or 1057, when the scribes in *Rus'* started utilizing the Cyrillic alphabet, which was purposefully designed for phonology of Old Bulgarian by the disciples of Cyril and Methodius, to copy and translate liturgical texts into their own East Slavic (or Old Russian) dialect. The Russian scribes omitted several characters which were redundant, however, Matthews points out that the Old East Slavic scribes had a proclivity for using abbreviations, which also included the use of titlos to form titlo-abbreviations.

We should keep in mind, that Gammanovich's and Sazonova's description of titlos above concerns the use of titlos in (Old) Church Slavonic, and that titlo-abbreviations in Old East Slavic is the locus of this thesis.

### 3.2.1 Handwriting

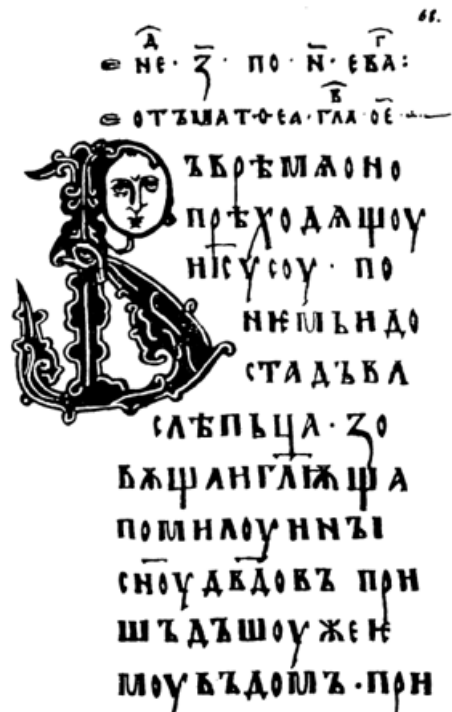
Tikhomirov summarizes the evolution of Old East Slavic handwriting in *Развитие русского кирилловского письма* (1966), based on В. Н Щепкин's *Учебник русской палеографии* of 1918. Old East Slavic handwriting can largely be divided into three consecutive scripts with their own characteristics: *ustav*, *poluustav* and *skoropis'* (and a transitory script between *poluustav* and *skoropis'*, sometimes referred to as *beglyj poluustav*).

### 3.2.1.1 Ustav

The oldest Russian codified handwriting was characterized by geometric lettering along a line, where the letters' upper and lower ends did not extend above the line or fall under it (Tikhomirov, 1966), c.f. Figure 5- Ustav (in the Ostromir Gospel) below.

Through the 11th and 14th century, the development of *ustav* the geometrical shapes changed, whereby the location of transverse crossbars in letters started slanting upwards in an oblique fashion. Letters also started to extend above and fall below the lines which the geometric lettering followed. For instance, in the older handwriting и was written like modern н, and н like the Latin letter N. Throughout the 12th century, this affected ю and љ, like its modern stylization. Ж lost its' symmetry, as the upper part decreased, as happened with с. The letter ч had its' «leg» or «tail» extended, and д fell further below the line, while ъ extended above it. Letters from the second half of the 12th century almost did not differ from the first half of the 13th century, however, in the second half of the 13th century handwriting saw abrupt change intertwined with the political upheaval brought upon *Rus'* by the Mongol invasion and the subsequent Tatar yoke (Tikhomirov, 1966).

Figure 5 – Ustav



(Gammanovich, 2001, p. 12)

From the second half of the 13th and during the 14th century, the geometrical features were abandoned, and handwriting was characterized by curvatures for a faster, more efficient, and simplified handwriting. The transition from *ustav* to *poluustav* was particularly visible in the loops of the letters у, ъ, ѣ, ѥ, and ѧ, which lost their geometricity. В and ж was distinguished by a smaller upper half, while the latter also had a antennae on top (Tikhomirov, 1966).

### 3.2.1.2 *Poluustav*

As manuscript designs diverged in between *Rus'* proper and the Grand Duchy of Lithuania, *poluustav* appeared (c.f. figure 6 below), and later also characteristic Russian and Belarusian-Ukrainian letters. *Poluustav*, apparently, saw its inception due to a need of clarity, rather than beauty, and was probably developed by scribes of business papers. Use of *poluustav* was particularly pronounced in documents in the 14th and 15th century. While still close to *ustav*, *poluustav* had several significant differences (Tikhomirov, 1966).

A new handwriting appeared in the 15th century due to south-Slavic migration to Russia, with some of its features being adopted by medieval Russian scribes, which led to a new type of *poluustav* particularly evident in ecclesiastical and literary writing, however, these features were much less evident in the 16th century. This style never permeated into business writings, due to the need of precise language in accordance with the spoken language (Tikhomirov, 1966).

Figure 6 - *Poluustav*



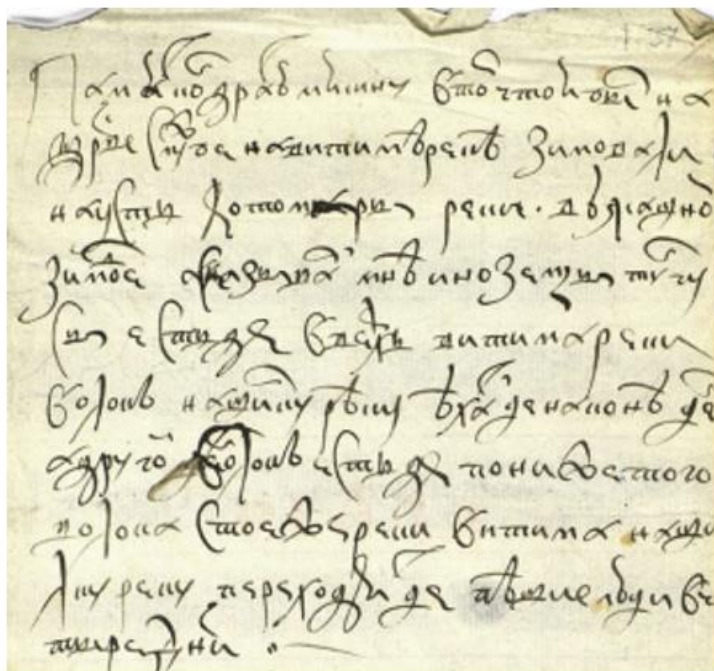


Throughout the 15th century, the outlines of what can be called a Muscovite *poluustav* developed, however, it was not limited to Muscovy (as it was also used in both Belarus and Ukraine), but it was typical for Muscovy. The characteristic of the Muscovite style *poluustav* was B with rounded outlines, Д that extended far below the line, and large curly 3 utilized, and the (similar) letter *zelo* (c.f. figure 7) was utilized in places it should not, before it merged with 3. By the 16th century, the three-legged τ was formalized, along with accents and partly division of words. The style of the 16th century Muscovite *poluustav* carried over into the 17th century, however, it was of much poorer fashion - attributed to the difficult economic situation. The *poluustav* of the 17th century has since been preserved in «Old Believer» manuscripts (Tikhomirov, 1966).

### **3.2.1.3 Skoropis'**

Throughout the 14th and 15th century *skoropis'*, or cursive, appeared, and it was mainly developed in business papers. This cursive style was based on *poluustav*. In general, early *skoropis'* did not differ much from *beglyj poluustav*. Every letter would be written separately, but they were characterized by a freer hand and sweeping motions. Particularly, *lettered titlos* were characteristic of the time (c.f. Gammanovich), which can be explained both by the accelerated speed of writing, but also due to the lack of writing materials, such as parchment and paper. *Skoropis'* became widespread in the 16th century when it was utilized in business papers and clerical writings, and private correspondence, which lasted up until the 18th century. *Skoropis'* was distinguished from *ustav* and *poluustav* in the sense that letters lost their graphical distinctiveness. Furthermore, letters became hooked, lost proportionality and letters became connected, along with bigger letters extending and protruding above and below the line. With its formalization in the 17th century, the individual character of cursive writing manifested – and reading cursive became a skill in its own (Tikhomirov, 1966).

Figure 8 - Skoropis'



(Архивы ргада о народах Забайкалья и реки Шилка, 2022)

#### 3.2.1.4 Titlos in *poluustav* and *skoropis'*

Titlos were frequently used in *poluustav*, like in *skoropis'*. However, words were written *pod titlom* in *poluustav*. That is, they used the *simple titlo* to a greater extent. *Skoropis'* also featured abbreviation of frequent words, but most commonly in the form of *lettered titlos*. Overall, the use of titlo-abbreviations was common in *poluustav* in the form of simple titlos, but titlo-abbreviations became far more common in the form of *lettered titlos* with the development of *skoropis'* (Tikhomirov, 1966).

Both *simple* and *lettered* titlo-abbreviations were utilized due to an economic incentive in terms of scarcity of materials, but also to streamline penmanship in terms of efficiency (with regards to both writing and reading), according to Tikhomirov. I stress, however, that this thesis concerns itself with *simple titlos*, which are simply addresses as titlos, unless otherwise stated.

#### 3.2.1.5 *Grazhdanskyj shrift*

The 18th century saw the introduction of the *grazhdanskyj shrift* along with the introduction of Arabic numeric characters, the variety in letter styles decreased throughout the 18th century, which heralded the inception of the modern style of Russian handwriting (Tikhomirov, 1966).

### 3.2.2 Writing materials

In *Slavic Paleography and Early Russian Printing: The Genesis Of The Russian Book* (1967), Kaldor describes the history of Slavic Paleography in *Rus'* and Medieval Russia.

According to Kaldor, it is not possible to establish that there was regular use of so-called primitive writing materials, like leaves, stones, metals, clay tablets, or papyrus in ancient *Rus'*. However, findings from excavations do suggest that waxed tablets were used (mostly likely to practice writing) from the 12<sup>th</sup> and throughout the 14<sup>th</sup> century in medieval Russian cultural and economic centers, such as Novgorod.

Birch-bark letters suggest that several cultural strata of Novgorod utilized bark, upon which letters were etched, to compile private documents from the 11<sup>th</sup> and throughout the 14<sup>th</sup> century. Kaldor states that there have been vague references to the use of birch-bark as a writing surface in ancient *Rus'*, however, the first material proof was found in Novgorod in 1951. Since its humble beginnings, the number of birch-bark documents has risen to thousands, not only in Novgorod, but also smaller quantities in Pskov, Rjazan', Smolensk and other places, which testifies to the scope of birch bark writing in medieval Russia (Kaldor, 1967; Nessel, 2015).

Beyond archeological, paleographical, and linguistic value, these birch-bark manuscripts reveal the scope of literacy among different classes of people in the ancient cultural and trade center of Great(er) Novgorod. The birch-bark letters range from children's drawings, puzzles and, word games to teaching tools for school children (alphabets, lists of numbers), letters (written by citizens, craftsmen, and merchants), *I owe you* notes, birthday greetings, trade contracts, bills of sale and invoices. The average birch-bark scroll measures 8 to 10 inches in length, and 2 to 3 inches in width, which accommodates for three to six lines of text (from 200 to 250 characters). This standardized size was attained by trimming the bark fragment on both ends. Relatively crude *ustav* letters were etched on the smooth inner surface of the birch-bark with a bone, wooden or metallic instrument, (probably) called a *pisalo* (a stylus [Nessel, 2015]). Whereas the use of ink on birch-bark was incredibly rare in medieval Russia, the use of bark as a writing surface remained prevalent from the 16<sup>th</sup> until the late 18<sup>th</sup> century, due to lack of actual paper to write on (e.g., in remote Siberia), even when ink was available (Kaldor, 1967).

Parchment was the standard writing surface for codices and official documents in medieval Russia and it was likely imported from either Greek or German Hanseatic merchants due to the fine qualities of the pages in 11<sup>th</sup> century Russian codices, which is substantiated by references in codices themselves, along with the continued use of birch-bark as a writing surface for many centuries. However, Russian parchment (of noticeably lower quality) was made primarily from calf skin, thereby the ancient Russian name *teljatina* (something made of calves, as *telja* means calf) for documents written on parchment, along with *kozha* (skin), *mekh* (hide or fur), *khartija* or *kharatija* (charta). The term *pergamen(t)* was introduced to Russian in more recent time, as a borrowing from either western Russian or Polish. The procedural treatment applied to the skins by parchment-makers was almost identical with the (ancient) process of both Greeks and Romans. Though, in contrast to South Slavic practice, Russian writers would not *palimpsest*, that is to superimpose new text on a repurposed writing surface (at least routinely), or mix different writing surfaces in codices (Kaldor, 1967).

The first documented use of paper for writing in medieval Russia is dated to the mid-14<sup>th</sup> century, while the oldest extant document written on paper is the (undated) charter of Prince Vasili Davidovich of Nizhnij Novgorod (d. 1345). The oldest extant Russian codex written on paper is a volume of the Sermons of Isak Sirin (d. 1381). The transition from parchment to paper was long and gradual, but parchment became reserved for ceremonial texts only in the late sixteenth century. The Russians unsuccessfully attempted to establish their own paper industry several times. Despite the existence of several paper mills over the course of centuries, the Russians largely resorted to import paper in large quantities from abroad (Kaldor, 1967).

### **3.3 Standardization of Russian**

In *The Palgrave Handbook of Economics and Language* (2016), Ginsburg and Weber outline the standardization of Russian, which occurred from the 16th century and onwards, as the individual principalities of *Rus'* were unified under Moscow. The state played a key role in standardizing of the Russian (language). This standardization occurred both with regards to reducing dialectical barriers amongst (ethnic) Russians, but also to expand to use of Russian in public affairs. The standardization of Russian was also a testament to how policies were formed based on the struggle between central and local elites, vying for economic and political control.

The individual principalities of *Rus'* were fragmented and isolated both during and after the Mongol yoke, up until the consolidation of power by the hand of the Muscovites throughout the 16<sup>th</sup> and 17<sup>th</sup> century. The fragmentation and isolation affected the (ancient) Russian language, which had diverged into various dialects, which in turn itself posed a challenge for the Muscovite state bureaucracy. As the centralized system of government developed, the need for a standardized language thus intensified. Religious literature was predominant in medieval *Rus'*, and as the church carried out schooling, the populace was taught Church Slavonic [CS]. CS was, however, based on southern Slavic dialects and thus different from both Russian and East Slavic in general. CS was much too archaic and complicated to be imposed as the standardized form. As such, the (so-called) Moscow official (or chancery) language – a synthesis of the many dialects of ancient Russian – formed the basis of which standardized Russian was codified.

The standardization of Russian language was a practical matter, as dialectical and linguistic diversity complicated communication in-between the centralized and local powers. As such the (so-called) Moscow official (or chancery) language became the foundations of standardized Russian due to Moscow's dominant role while unifying the Russian lands from the 16<sup>th</sup> century and onwards, however not through evolution, but through purposeful implementation of a state policy.

## 4 Theory

### 4.1 Tsar & God

In *The Sacralization of Monarch in the Context of Historical and Cultural Development* (2012) Zhivov & Uspenskij describe how the Muscovite monarch assumed a religious character through orthographical means, which led to a juxtaposition between God and the Muscovite monarch. The parallelism between the monarch and God and their shared attributes came to *Rus'* as early as the 6<sup>th</sup> century through the works of *Agapit* (Agapetos), who stated that while the tsar (i.e., the monarch) was perishable like any man, he was also like God through his power. However, it was explicitly stated that the monarch's power was not autonomous, but subject to God's moral law, as it was ordained by God. A corollary to this was that the will of a lawful tsar was the will of God (however with certain caveats). The notion inherited from Agepetos' work was reiterated in the Hypatian Chronicle, amongst others, where it was written: *Although the tsar's earthly nature is like that of every man, the power of his rank is higher, like God* (Zhivov & Uspenskij, 2012, p. 4). From Vasily II the

Blind's rule and henceforth the Muscovite ruler was regularly referred to as tsar (derived from the Latin title Caesar [царь, 2023]), and his great grandson Ivan IV the Terrible was coronated as tsar through religious rite. With time, Muscovite scribes were also instructed that the title of the monarch, namely tsar, was to be written *pod titlom*, that is, under a title. According to the old tradition, tsar was written *pod titlom* only when it referred to God (as The Heavenly Tsar), whereas in reference to the muscovite tsar, it was to be written out letter by letter and not *pod titlom*, even if the given tsar was considered holy. With time however, grammatical works instructed the scribes to write the names of pious tsars *pod titlom*, while it was to be written letter by letter while referring to unlawful tsars. Consequently, the attributes and divinity of God (the Heavenly Tsar) were extrapolated to the Muscovite ruler (the earthly tsar), and thus the tsar was embedded in the religious tradition. The Muscovite's perception of the ruler and his title's sacred or divine nature was also attested by European travelers' writings (c1700th century), who even claimed that the Muscovites thought that the title tsar itself was created by God, and not by man.

#### **4.1.1 Semiology**

Zhivov & Uspenskij's work is written within the theoretical perspective of semiology (or semiotics). While I will not delve deep into the subject matter of semiology, a brief introduction is in order. The crux of Semiology is further very much compatible with Cognitive Linguistics which I will explicate upon in section 4.4.

Saussure, the progenitor of Semiology, is considered the father of modern linguistics, and in his view, words and languages are to be viewed as the collective product of social interaction, which provides for the conceptual framework by which we analyze and vocalize the description of reality. Furthermore, language is a whole system, which is not built by constituent parts that exist independently, but as a system where the smallest part does not exist without the whole (Saussure & Harris, 2021).

#### **4.2 Diglossia**

Uspenskij presented his theory of diglossia in *Rus'* in several works, e.g., in *Istorija russkogo literaturnogo jazyka* (1987). However, I will not delve into Uspenskij's own works on diglossia in this thesis, as the relevant content of the theory is summarized in the critiques raised by Grannes in his *Review of История русского литературного языка (XI-XVII вв.)* (1989), Lunt in *History, Nationalism, and the Written Language of Early Rus'* (1990), and Collins in *On Diglossia and the Linguistic Norms of Medieval Russian writing* (1992), which

I will present below. The reasoning for addressing this theory (and the debate on its applicability) is that it may shed light on how (Old) Church Slavonic norms and literary practices affected Old East Slavic scribes and their practices.

As stated by Grannes (1998), Uspenskij utilizes Ch. A. Ferguson's description of the phenomenon, which he described in *Diglossia* (1959). According to Ferguson, diglossia describes a situation, in which two variants of the same language coexist in a speech-community. This phenomenon is similar to both dialects and bilingualism, but diglossia, however, dichotomizes between a *high* and a *low* variant of the same language where both iterations adhere to codified norms, where the respective variants are utilized in different situations.

In Uspenskij's model of diglossia in *Rus'*, Slavonic represents the *high* variant while it was also *the* literary language of medieval *Rus'*, acquired through formal study, whereas East Slavic was the *low* variant and the vernacular language (although not to be confused with the colloquial and spoken language in this context). The proponents of diglossia, such as Uspenskij, stress the role of linguistic consciousness among scribes in medieval *Rus'*, which allowed the scribes to utilize linguistic variation according to separate sets of static and codified learned norms, evident by the features which are realized in the written tradition (Collins, 1992; Lunt, 1990).

According to Uspenskij, Slavonic supposedly acquired normative status in *Rus'* as it was the language which all Orthodox believers had to know as a part of religious education, and as Slavonic was the language of divine truth – it was reserved for writings which concerned itself with the divine and higher objective truths (i.e., it adhered to canonical Orthodox ontology), and not ordinary conversation and subjective impressions, which was the realm of East Slavic. Consequently, the writer's attitude to the theme would also dictate whether he utilized Slavonic or the vernacular in text. A switch in code or norms in text, from *high* to *low* and vice versa would thus also occur depending on the authors perception of objectivity and subjectivity to the subject matter. Furthermore, proficiency in Slavonic was supposedly acquired passively, through reading and memorization, as the ability to write orthographically must have required special training (Collins, 1992; Grannes, 1989.).

In Uspenskij's view, (Church) Slavonic, which was inherited from the South Slavs, was adapted into Russian (Church) Slavonic, which later fulfilled the role as the literary language

of the East Slavs, who reputedly viewed it as the written, learned form of their own vernacular tongue. According to Uspenskij's theory, the diglossic situation of Slavonic and East Slavic in complementary distribution developed already in the 10<sup>th</sup> or 11<sup>th</sup> century and persisted well into the 17<sup>th</sup> century, whereas the transition to bilingualism occurred in the late 17<sup>th</sup> century (Grannes, 1989).

Uspenskij further states that the period of diglossia was characterized by three separate waves of South Slavic influence: 1<sup>st</sup> when Kiev adopted Byzantine Christianity and the literary tradition of the Bulgaro-Macedonian church, 2<sup>nd</sup> when a significant change in linguistic consciousness occurred, whereby Slavonic was perceived as an independent norm rather than as a codified form of the contemporary spoken language at the end of the 14<sup>th</sup> century, and a 3<sup>rd</sup> wave in the 17<sup>th</sup> century (Collins, 1992).

In terms of the 2<sup>nd</sup> wave of South Slavic influences, Uspenskij raises an interesting issue relevant to this thesis. According to Uspenskij, muscovite scribes wanted to differentiate between homonyms and polysemes orthographically to justify phonetic variation (during the 16<sup>th</sup> and 17<sup>th</sup> century), c.f.:

*(...) they pronounced аγγελъ with titlo ('angel') differently from the same spelling without a titlo ('fallen angel') and assigned different meanings to o-stem slovo, slova ('word') and consonant-stem Slovo, Slovese ('the Word'). Similarly, certain proper names were distinguished by their stress: Marija 'the Virgin Mary', but Márija/Már'ja for other saints or as a baptismal name; Sofija 'Holy Wisdom', but as baptismal name Sófija/Sóf'ja (...)* (Collins, 1992, p. 82-83).

What is interesting here is the apparent attitude towards titlos – phonetic variation would also imply that inherent meaning of titlo-abbreviated words manifested in a conscious manner, and that the writer wished to raise certain notions with the reader, or as described in Collins critique: *(...) as though the signans were related to the signatum in a nonarbitrary manner (...)* (Collins, 1992, p. 83). In terms of semiotics (or semiology) *signans and signatum* refers to the dual entity of an orthographic sign, which cannot exist without the other – that is, that it holds meaning to both the writer and the reader (Saussure & Harris, 2021).

Lunt underlines that the language in East Slavic manuscripts ever since the Ostromir Gospel of 1056-57 testifies to a linguistic dualism, which some scholars have described as a Russian recension of Old Church Slavonic (or Old Bulgarian). One theory interprets this as a process



of russification, while another, such as the one presented by Uspenskij, hold that the East Slavs had their own native literary language. Lunt however, states that ([Old] Church) Slavonic and East Slavic were variants of one language, but without a codified system of contrasting *high* and *low* features, or the characteristics of diglossia or bilingualism.

Grannes reiterates D.S. Worth's critique from 1978, where he asserted that there was only diglossia in the early period: [...] *medieval Russia conforms to the diglossic criteria at most for one-fourth of the period from the 10th to the 17th centuries, and then (14-16cc.) only partly* (Grannes, 1989, p. 261-262), which is substantiated by Lunt, who's also cited by Grannes: *in the 10th and 11th centuries there was still only one single Slavic language*", and *"and therefore neither diglossia nor bilingualism* (Grannes, 1989, p. 262). Lunt does state that regional differences appeared by the 13<sup>th</sup> century, although the written language remained stable and in accordance with norms. Despite south Slavic influences, the written language was still very close to the spoken vernacular. However, Lunt says that diglossia may have been emerging in the 13<sup>th</sup> century (when Uspenskij claims that bilingualism was manifesting). Collins, on the other hand, concludes that Uspenskij's model may be applicable to later Muscovite writings, but not to earlier writings due to the absence of evidence of a linguistic norm - available evidence does not support the idea that Russian scribes had a clear conscious conception of a literary norm. Collins points out, that the definition of diglossia presumes that there was a linguistic consciousness evident by the scribes' use of language, however, the available evidence does not testify to this. Lunt also stresses that the evidence does not support Uspenskij's theory of diglossia in *Rus'*.

### 4.3 Cognitive Linguistics

Ungerer & Schmidt's concise explanation of cognitive linguistics presented in *An Introduction to Cognitive Linguistics* (2006) makes for an ideal starting point before going in depth on the more complex aspects of cognitive linguistics. In essence, cognitive linguistics ask what goes on in the mind of speakers when they give meaning to words and sentences? Furthermore, cognitive linguistics also ask how associations and impressions drawn from experience give meaning to words and sentences, how we interact with the world and how we interpret prototypes and categories, which are not static, but shifting according to context, by cognitive and cultural models stored in the mind.

Thus, we can delve deeper into the subject matter, and the *Handbook of Cognitive Linguistics* (2015) by Dabrowska & Divjak is our springboard.

Based on the assumption that language is embedded in our general cognitive abilities, and that it is not a separate faculty, cognitive linguistics claim that meaning is conceptualization, and the grammar is shaped by language use. An ever-increasing amount of linguistics subscribe to cognitive linguistics, who are all united around the cognitive commitment, as described by Lakoff (reiterated by Dabrowska & Divjak, 2015, p. 1):

*All cognitive linguists are, or should be, committed to providing a characterization of the general principles of language that is informed by and accords with what is known about the mind and brain from other disciplines. It is this commitment that makes cognitive linguistics cognitive, and thus an approach which is fundamentally interdisciplinary in nature.*

Several other assumptions follow the cognitive commitment. The first being that cognitive linguists share a usage-based view on grammar and that grammar is shaped by language use. We must research how it is used to understand how a language is learned and structured. Second to this, cognitive linguistics assume that general cognitive faculties, such as perception, attention, memory, categorization, and abstraction structure and interpret language systems through use. Third, the purpose of language according to cognitive linguists is to convey meaning, and as such every facet of language is meaningful, including grammatical constructions. The term *meaning*, in turn, is based on conceptualization, embodiment, dictionary and encyclopedic information (Dabrowska & Divjak, 2015).

### **4.3.1 The Cornerstones of Cognitive Linguistics**

In *Language Change and Cognitive Linguistics: Case Studies from the History of Russian* (2022), Nessel details the four commitments that make up the cornerstones of cognitive linguistics, namely 1) the cognitive commitment, 2) the semiotic commitment, 3) the network commitment, and 4) the usage-based commitment, which is summarized below.

#### **The cognitive commitment**

The most important prescription of Cognitive Linguistics is the cognitive commitment, which describes that domain-general cognitive processes shape language use, which means that language is not separate from other cognitive abilities, such as perception. Nessel reiterates research by Fedorenko and Shain, who assert that while language is specialized to certain areas of the brain, the mental processes surrounding language are not different from other faculties such as perception or cognition. As such, the cognitive commitment is an ontological

assumption, followed by a methodological imperative that states that inquiry into a linguistic phenomenon must be related to general, or other cognitive faculties and their processes (Nesset, 2022).

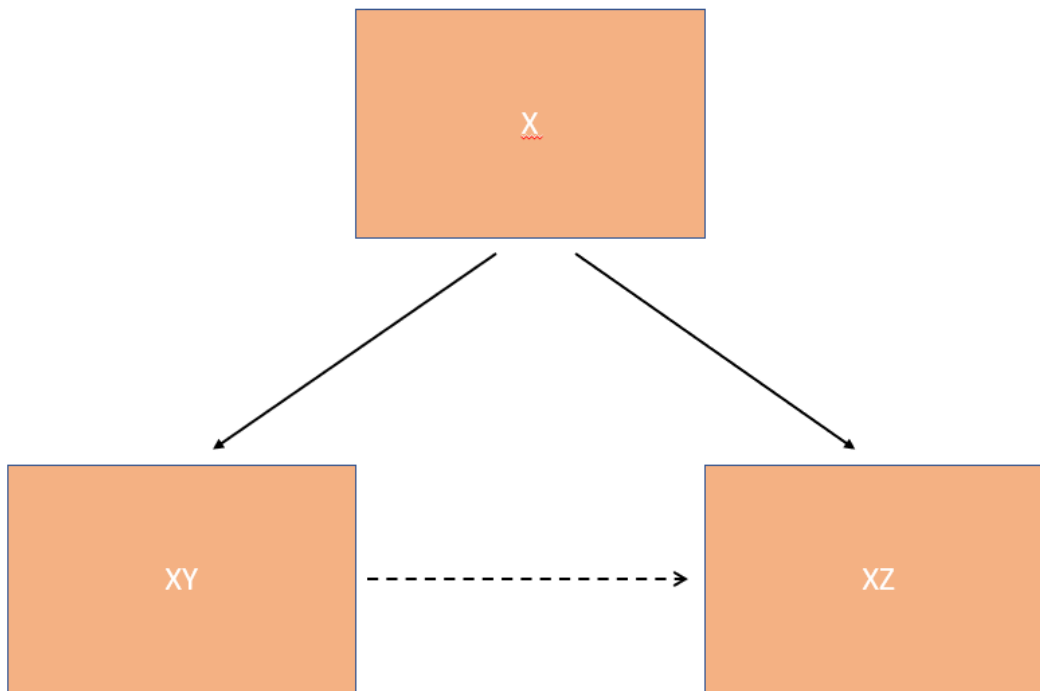
### **The semiotic commitment**

The semiotic commitment says the prime purpose of language is to convey meaning, which is hardly a radical position presented by cognitive linguists, however, cognitive linguists thus analyze language in terms of representations that connect form and meaning *only*. The only grammatical structures that can exist in the scope of cognitive linguistics is 1) phonological, semantic, or symbolic structures which are observed in linguistic expressions, e.g., sounds or handshapes, 2) schemas for these structures - that is - semantic or pragmatic meaning in a broader sense and 3) categorizing relationships in between 1) and 2), e.g., morphological patterns and syntactic constructions (Nesset, 2022).

### **The network commitment**

The network commitment presupposes that all linguistic categories form networks of related subcategories that are related in one form or another. These relations are motivated by a prototypical subcategory, and non-prototypical subcategories which are related to the prototype in one way, and sometimes to each other in another way – they may have shared similarities with the prototype, but also differences (Nesset, 2022).

Figure 9 – The Network Commitment



(Based on *Language Change and Cognitive Linguistics: Case Studies from the History of Russian*, Nessel, 2022)

Furthermore, in cognitive linguistics, language is perceived as being based on constructions, which were (briefly) addressed in section 1.3. (c.f. Janda & Dickey, 2017).

### **The usage-based commitment**

In Cognitive Linguistics, grammar is viewed as conceptualized generalizations that emerge bottom-up – grammar is not prescribed from a lexicon but emerge through language use. There is no innate ability to construct grammatically correct sentences, grammar is a by-product and shaped by language use (Nessel, 2022).

The usage-based approach describes Cognitive Linguistics ‘s applicability to statistical methods, as (electronic) corpora provide vast amounts of authentic linguistic examples and their usage, and thus allows for extrapolation of data to general rules based on observed use (Nessel, 2022).

### 4.3.2 Metonymy and metaphor

As described by Haser (*Metaphor, Metonymy, and Experientialist Philosophy: Challenging Cognitive Semantics*, 2005), a major issue of studies into metaphors has been inconsistency in the delineation of the subject matters. Some researchers have utilized broad definitions for metaphors, which have blurred the line in between metaphor and related concepts of figurative meaning, while others have utilized narrower definitions reducing metaphors to extension based on similarities. The latter, narrow definition of metaphors is oftentimes employed by cognitive linguists, however, it has still proved to be hard, or even impossible to adequately differentiate between metaphor and metonymy, and alas, there is not necessarily consensus among cognitive linguists.

The difference between metaphor and metonymy is often described as similarity and contiguity, respectively, where metonymies refer to contiguity between a source and target, and metaphors (or metaphorical mappings) seem to be motivated by similarity – not between source and target – but similarity and analogy between dissimilar concepts. Accordingly, metaphors are thought to describe mapping across different domains, in contrast to metonymy (or metonymic extensions) which takes place within one domain. Curiously, detailed descriptions of the difference between metaphor and metonymy are most often found in literature concerning metonymy, more so than literature on metaphors (Haser, 2005).

### 4.3.3 Defining Metonymy

Brdar (*Metonymy and Word-Formation: Their Interactions and Complementation*, 2017) like Haser, also states that there's not necessarily a consensus view as to defining metonymy. Brdar, however, reiterates Lakoff and Johnson's description of metonymy «as the use of *one entity to refer to another that is related to it*», and Köveses' and Radden's explication upon this, as: «a cognitive process in which one conceptual entity, the vehicle, provides mental access to another conceptual entity, the target, within the same domain, or [Idealized Cognitive Model]» (Brdar, 2017, p. 30). In essence, metonymy provides a mental shortcut by which we may refer to entities that have no current or convenient linguistic expression.

Like Haser, Brdar also explicates upon how metonymy can be defined by juxtaposing metonymy with metaphors. Whereas metaphors have been scrutinized with disregard to metonymic properties – metonymy has oftentimes been subsumed under metaphors as an instantaneous form. Furthermore, metaphor has proven to be a much more common topic in linguistic research, than metonymy.

Brdar, further describes the three central differences between metonymy and metaphors, which are 1) that metonymy is based on contiguity, whereas metaphor is based on similarity. 2) That metonymy and metaphor differ in the number of conceptual domains involved and 3) that metonymy and metaphor differ in directionality of the involved conceptual mappings.

These three central differences are explicated upon below:

The first difference is that metonymy, or contiguity cover all associated relations except for similarity, which is covered by metaphor. Metaphor oftentimes involves two or more entities that exhibit similarities, however without an explicit statement of this – you explain a concept by addressing another concept which implicitly exhibit or share qualities or features with the invoked concept (Brdar, 2017).

The second major difference between metonymy and metaphor involves the number of distinct conceptual domains we are dealing with. Metonymy occurs within a single domain, whereas metaphor occurs two (or more) domains (Brdar, 2017).

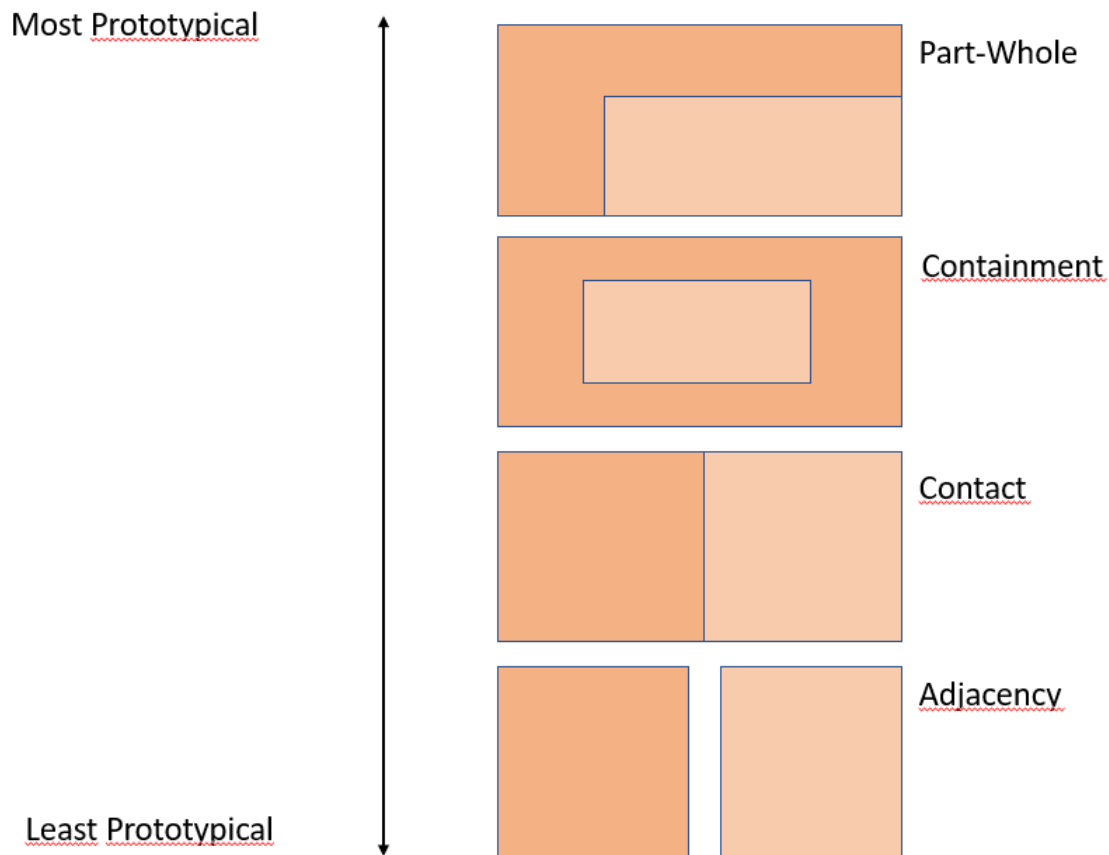
Third, the directionality of metonymy and metaphors differs. Metaphor is most often unidirectional – a tangible concept and domain is utilized to explain a social or mental domain. Metonymies on the other hand, may occur in any direction, from the tangible to the abstract and vice versa (Brdar, 2017).

At the core of metonymy is the part-whole relationship, which is invoked by an expression (or metonymic vehicle) which is associated with a metonymic source within a conceptual domain. It is clear, that metonymy is an intra-domain phenomenon, so that we do not concern ourselves with identifying domains and subdomains and the jump from one domain to another while dealing with metonymy – no domain mapping occurs at all (Brdar, 2017).

#### **4.3.3.1 Typologies of metonymical relations**

The intra-domain nature of metonymy and contiguity relations is visualized by Nessel in *“Cyclic” time in the history of Russian: Culture and language internal factors* (2016, p. 67), where he renders Peirsman’s and Geeraerts’ typology of metonymical relations:

Figure 10 - Typology of metonymical relations



#### 4.3.3.2 Abbreviations

In the literature some writers hold that abbreviations, such as alphabetisms, acronyms and other types of shortening of words involves conceptual metonymy, and that the link in between the full word and the abbreviated form is metonymic in nature. However, metonymy is dependent on an appropriate context to trigger a shift, and it is often hard to discern if there is an appropriate context in which alphabetisms and acronyms occur (Brdar, 2017).

#### 4.3.4 Polysemy

Polysemy is most often defined as a form ambiguity where two or more related senses are associated with the same word, as in: *consider the meanings of glass in “I emptied the glass” (‘container’) and “I drank a glass” (‘contents of the container’)* (Dabrowska & Divjak, 2015, p. 472).

In the scope of cognitive linguistics, polysemy is treated as (i) *viewing meaning/sense as categorization*, (ii) *recognizing the importance of context for meaning/senses and that*

*linguistic and encyclopedic knowledge are hard to keep separate, and (iii) incorporating prototype theory into linguistics* (Dabrowska & Divjak, 2015, p. 473). In terms of (i), lexical items are coded according to the mental categories we have established to order concepts. In terms of (ii) the meaning of a lexical items is hard to make out without context or encyclopedic knowledge of the real world, in which you frame it. Finally, in terms of (iii), prototypes, are those concepts with high que validity, that is, i.e., they are more representative for a category, than other concepts which still have features which put them in the same category. Now, prototypes also lead to radial categories, which is a network of categories that all share some features, while they are also different. The proximity between categories is determined similarities and differences (Dabrowska & Divjak, 2015).

## 5 Research Question(s)

We have learned from Gammanovich and Sazonova (see section 2), that titlos indexed a specific, sacred meaning in (Old) Church Slavonic, and that abbreviation by titlo was reserved for particularly sacred or revered objects, though, this thesis concerns itself with Old East Slavic. However, similar use of titlos in described by Zhivov & Uspenskij in Muscovy (see section 4.1). Moreover, we have learned from Collins (see section 4.2), that the Muscovites would pronounce titlo-abbreviated nouns different from those not titlo-abbreviated.

As described by Matthews (see section 3.2) the Muscovites had a liking for utilizing titlo-abbreviations, and according to Tikhomirov (see section 3.2.1) changes in script promoted the gradual increase of use of titlo-abbreviations in Old East Slavic, while we have also learned from Kaldor, that writing materials were scarce and expensive in *Rus'* (see section 3.2.2). These perspectives imply pragmatic use of titlo-abbreviations.

Arguably, this leaves us with two hypotheses. One stating that titlo-abbreviations were used purposefully to convey meaning in Old East Slavic, like in (Old) Church Slavonic, c.f. Zhivov & Uspenskij, and an alternative hypothesis, that states that titlo-abbreviations were used pragmatically due to economic incentives and in terms of efficiency, c.f. Tikhomirov and Kaldor.

Now, let us look back to section 1.2, where I stated the simplified research question: *why did Old East Slavic scribes use titlo-abbreviations?* Now, we have two hypotheses drawn from the literature, which may answer this question: 1) that titlo-abbreviations were used to express *meaning* and 2) that titlo-abbreviations were used pragmatically. However, the first



hypothesis also raises the corollary question: What is the *meaning* of titlo-abbreviations in Old East Slavic? It is probably no surprise that my belief is that the first hypothesis answers the research question by reading into the thesis' framing, but the nature of the titlos' meaning remains somewhat unclear to me. However, this thesis attempts to answer both research questions.

## **6 Data**

### **6.1 Digitalization and electronic corpora**

As presented by Nessel in *When We Went Digital and Seven Other Stories about Slavic Historical Linguistics in the 21st Century* (2017), scarcity of data has been a pitfall within historical linguistics. However, the recent development of historical corpora of Slavic languages, such as 1) the Russian National Corpus, 2) the Czech National Corpus, 3) the PROIEL corpus of Old-Indo European languages including OCS, 4) the TOROT (Tromsø Old Russian and OCS Treebank), and 5) the Manuskript corpus of Slavic and Russian texts and 6) the RRuDi Corpus (Regensburg Russian Diachronic Corpus) have mended the issue.

These corpora resolve the issue with scarcity of data given they are 1) large and 2) have good annotations, and Nessel assures that both the PROIEL and TOROT corpora are representative examples in terms of size. However, in comparison to the Russian National Corpus, these are still small corpora. Though, both PROIEL and TOROT is equipped to detect robust tendencies for the more frequent linguistic patterns, but not for infrequent phenomena. Furthermore, both PROIEL and TOROT are lemmatized, with part of speech and morphological annotation, which makes it possible to search for all forms of a particular lexeme and specify searches for inflectional categories. Both corpora feature syntactic annotation, which enables inquiry into specific syntactic constructions (Nessel, 2017).

### **6.2 The Tromsø Old Russian and OCS Treebank**

The Tromsø Old Russian and OCS Treebank (TOROT) described by Nessel is further explicated upon by Eckhoff & Berdicevskis (2015) in *Linguistics vs. digital editions: The Tromsø Old Russian and OCS Treebank*. At the time of publishing in 2015, the corpus contained approximately 160000 word tokens of Old Church Slavonic based upon Codex Marianus and Codex Supraliensis, 85000 word tokens Kiev-era Old East Slavic, and 60000 word tokens 15th to 17th century Middle Russian, all fully annotated.

Since then, the treebank has been expanded in two ways: 1) the developers have added additional layers of annotation and 2) the developers have added both Old East Slavic and Middle Russian texts, along with a modern Russian stage, to enable comparison of historical data with modern data.

Eckhoff & Berdicevskis state the key strength of TOROT is that it was designed and made for linguists, by linguists. The annotations are recognizable from a traditional point of view, while being based on contemporary linguistic theory. The TOROT treebank provides both syntactic analyses of every sentence, which is advantageous for both part-of-speech assignment and morphological analyses (Eckhoff & Berdicevskis, 2015).

### 6.2.1 The data in this thesis.

The data utilized in this thesis was compiled from The Tromsø Old Russian and OCS Treebank (TOROT) by Hanne Eckhoff.

The first dataset was compiled of all words forms shorter than their lemma, except for the verb *быти* (to be), pronouns and prepositions. These observations were controlled for whether they were abbreviated with a titlo or not. Upon closer inspection of this dataset, a frequency threshold of 10 or more observations of titlo-abbreviations was applied for nouns, to compile a second dataset. The relevant nouns which make up this second dataset is listed in the table below (Skjølsvold, 2021):

*Table 1 - Nouns*

<b>NOUN</b>	<b>TRANSLATION</b>
ЧЕЛОВѢКЪ	human
АВЪГУСТЪ	august
ГРИВНА	grivna
АРХИЕПИСКОПЪ	archbishop
ГЛАГОЛЬ	word
БОГОРОДИЦА	mother of God

АПОСТОЛЪ	apostle
ГОСПОДИНЪ	lord
ГОСПОДЪ	lord
ГОСУДАРЪ	sovereign
МЪСЯЦЪ	month
МУЧЕНИКЪ	martyr
НЕДЪЛЯ	week
СВЯТИТЕЛЪ	saint
СЪРДЪЦЕ	heart
ХРЪСТОЛЮБЪЦЪ	Christ-lover
ЦЪСАРИЦА	tsarina
ЦЪСАРЪ	tsar
ЦЪСАРЪСТВО	tsardom
ЦЪСАРЪСТВО	tsardom
ЕПИСКОПЪ	bishop

The second dataset, which is analyzed in this thesis, is composed of all iterations of the nouns above the frequency threshold. Note, that this also includes several abbreviations without titlos. These observations were organized in a data frame listing source (name of the original text), citation (reference, if it exists), ID (token number in the corpus), sentence ID (sentence number in the corpus), lemma (dictionary form), titlo (whether the word has one or not), additional grammatical information, lemma length (number of letters), form length (number of letters), length difference (number of letters) and context (the sentence the token occurs in) for all observations. The total amount of observations amounted to 3387 (Skjølsvold, 2021).

## 7 Methodology

### 7.1 Best Practice

In *Cognitive Linguistics: The Quantitative Turn*, Janda (2013) highlights some of the achievements of cognitive linguistics, such the implementation of data analysis within a theoretical framework, while addressing key points that should be regarded in terms of best research practice. Janda underlines of importance of public archiving of both data and code, even though transparency does not guarantee integrity. It does however go a long way in terms of uncovering fraudulent or mischievous intentions.

While Cognitive linguistics is at the forefront in terms of implementing data analysis within a theoretical framework, Janda stresses the continued role of introspection in linguistics research, as over-reliance on introspection in favor of observation has its limitations and undermines the scientific aspect of linguistics. Even though introspection has its pitfalls, it is still relevant. Introspection is after all the source of inspiration for hypotheses, which in turn can be tested through quantitative means, and introspection further serves a function in interpretation of results.

### 7.2 Statistical models and tests

As described by Levshina (2015, p. 253): *Logistic regression models the relationships between a categorical response variable with two or more possible values and one or more explanatory variables, or predictors*. Thus, logistic regression is well equipped to predict the outcome between to dichotomous options, such as in this thesis, where the inquiry is whether a noun is titlo-abbreviated or not. In this case, logistic regressions modelling is called binomial or dichotomous.

The statistical analysis in thesis is compromised of a CART Analysis verified by Random Forest and a fitted General Logistic Regression model. Tree models (such as CART analysis and Random Forest) and regression is described in the sections below:

#### 7.2.1 Correlation & Regression

The term correlation describes the degree of relationship between two variables. With greater correlation, it is easier to predict the value of one variable if we know the value of the other variable(s). There are two caveats to correlation: 1) correlation coefficients assume that the relationship between variables is linear, even though there may be many other possibilities. 2)

Correlation does not mean that there is a causal relationship between the variables (Janda, 2013).

The line of best fit, or regression line, which is described by correlation is the basis for regression models and its equation is what predicts outcomes, based on the input, however, it is only a perfect fit when the correlation between variables is perfect. A model is rarely a perfect fit, and therefore there is often some difference between predicted and actual values. This difference is called «error», and the regression model's equation's fit is reported by its standard error of estimate, which consequently is a measure of how well the regression equation fits the data. Regression inherits the drawbacks of correlation due to its nature. (Janda, 2013).

Regression models are designed to predict a dependent variable based on one or more independent variables – which ideally should be independent of each other to avoid collinearity. In logistic regression, for example, the dependent variable only has two values, which is particularly useful for linguistic research such as in this thesis (as we have already learned from Levshina). Regressions also provide measures (p-values and  $r^2$ , along with index of concordance), which indicate the significance of the data sample, the variables, the quality of the model. To avoid overfitting, the regression model can be trimmed for insignificant variables. The respective models can be compared through chi-square test, ANOVA, or the Akaike Information Criterion (AIC) to see which one is best (Janda, 2013).

#### **7.2.1.1 T-test & ANOVA**

Analysis of Variance (ANOVA), based on the T-test, cannot be explained without understanding variance, which is *the shape of a distribution in terms of deviations from the mean* (Janda, 2013, p. 17). Half of the deviations from the mean will be positive, and the other half negative, thus the sum will always be zero. Variance is then measured as the sum of the squared deviations, divided by the number of scores in the given distribution, and the square root of the variance gives the standard deviation. Now, ANOVA divides the total variation among scores in two groups, the within groups variation where variance is the result of chance and a possible treatment effect. By putting the between groups variation in the numerator and within-groups variance in the denominator, we get the F-ratio. If  $F \leq 1$ , the variance is greater than or equal to the between groups variance – i.e., there is no treatment effect, but if  $F \geq 1$ , the higher values demonstrate a greater treatment effect and thus the ANOVA can give P-values which are indicative of significance (Janda, 2013).

## 7.2.2 Alternatives to Regression

Corpus data is often skewed, and not distributed along a bell curve (that is, normal distribution), which is assumed by regression models along with the assumption that all possible iterations of observations are represented in the dataset. Linguistic data often contains gaps, as all possible paradigmatic combinations are not observed. The problem with these assumptions, and the weakness of correlation, can be solved by 1) *classification and regression trees in combination with random forests*, and 2) *naive discriminative learning*. Tree, forest, and naive discriminative learning models also come with the advantage the researchers do not have to build the models themselves, and that these models also validate themselves, so that the researcher does not have to continue to collect additional data for validation (Janda, 2013).

## 7.3 Issues with Old Russian and R Markdown

Due to unresolved issues between Old Russian and R Markdown, the visuals (and prose) in the statistical analysis in this thesis is somewhat lacking. Both source labels and lemmas had to be relabeled (1-46, and a-v, respectively) to make the graphics in the statistical analysis readable, due to long source titles and lemmas that would not format properly. This comes with the major drawback that we are forced to consult a list of levels for both source and lemma (which is provided in the analysis) to interpret both graphics and summaries. I should also add that I was forced to toggle between system language(s) several times, due to diacritics in the source titles, while working with the data in frame R. All workarounds had to be done in R, as the .csv file containing the data of which the data frame was drawn would get corrupted and terminated when I attempted to save edits. Finally, R Markdown continued to refuse to knit the Old Russian letter *yat* (ѣ), which exclusively shows up as `<U+0463>` (Skjølvold, 2021).

## 7.4 R Markdown packages

The statistical analysis in this thesis was made possible by the use of the respective R (Markdown) packages made by Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. 2012; Bates, D., Maechler, M., Bolker, B., & Walker, S., 2015; Fox, J., & Weisberg, S., 2019; Harrell Jr., F. E. 2021a & 2021b; Hothorn, T., Hornik, K., & Zeileis, A., 2006; Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. 2006; Sarkar, 2008; Strobl, C., Boulesteix, A., -L., Zeileis, A., & Hothorn, T., 2007; Strobl, C., Boulesteix, A., -L., Kneib, T., Augustin, T., & Zeileis, A.

2008; Wickham et al., 2019; Wickham, H., Francois, R., Henry, L., & Muller, K. 2021; Wickham, H. 2016; Xie, Y., Allaire, J. J., & Grolemund, G. 2018; Xie, Y., Dervieux, C., & Riederer, E. 2020 (Skjølsvold, 2021).

## 8 Statistical Analysis

The following statistical analysis which spans chapter 8 was submitted for examination in the class HIF-3082 (*Titlos in Old Russian literature*, Skjølsvold, 2021) and was carried out with the intent of being utilized in this master thesis. What follows below is largely identical to the original paper, except for minor changes to text and formatting.

### 8.1 Levels

Following are the respective levels found in under Source and Lemma in a listed format. When consulting graphics and summaries, these lists should be consulted to interpret the results (Skjølsvold, 2021).

#### Source:

```
## [1] "Afanasij Nikitin's journey beyond three seas"
## [2] "Birch bark letters"
## [3] "Burtsov's alphabet"
## [4] "Charter of Prince Jurij Svjatoslavich of Smolensk on the alliance
with Poland and Lithuania, 1386"
## [5] "Colophon to Mstislav's Gospel book"
## [6] "Colophon to the Ostromir Codex"
## [7] "Correspondence of Peter the Great"
## [8] "Domostroj"
## [9] "Letter of E. Klementiev to F. M. Chelishchev"
## [10] "Letter of F. I. Golitsyna to V. V. Golitsyn"
## [11] "Letter of M. D. Kurakina to B. I. Kurakin"
## [12] "Letter of M. M. Shcherbatov to D. M. Shcherbatov"
```

- ## [13] "Letter of U. S. Pazukhina to S. I. Pazukhin"
- ## [14] "Letters of D. V. Mikhalkov to M. I. Mikhalkova and P. D. Mikhalkov"
- ## [15] "Letters of V. B. Golitsyn to Vl. B. Golitsyn"
- ## [16] "Life of Sergij of Radonezh"
- ## [17] "Life of Stefan of Perm"
- ## [18] "Materials for the history of the schism"
- ## [19] "Missive from Prince Ivan of Pskov, 1463-1465"
- ## [20] "Missive from the Archbishop of Riga to the Prince of Smolensk"
- ## [21] "Mstislav's letter"
- ## [22] "Novgorod service book marginalia"
- ## [23] "Novgorod's treaty with Grand Prince Jaroslav Jaroslavich, 1266"
- ## [24] "Russkaja pravda"
- ## [25] "Statute of Prince Vladimir"
- ## [26] "Testament of Ivan Jurievich Grjaznoj"
- ## [27] "The 1229 Treaty between Smolensk, Riga and Gotland (version A)"
- ## [28] "The First Novgorod Chronicle, Synodal manuscript"
- ## [29] "The Kiev Chronicle, Codex Hypatianus"
- ## [30] "The Life of Avvakum"
- ## [31] "The list of the Novgorodians' losses"
- ## [32] "The Primary Chronicle, Codex Hypatianus"
- ## [33] "The Primary Chronicle, Codex Hypatianus, PRE-PARSED"
- ## [34] "The Primary Chronicle, Codex Laurentianus"
- ## [35] "The Suzdal Chronicle, Codex Laurentianus"



```

## [36] "The Suzdal Chronicle, Codex Laurentianus, PRE-PARSED"

## [37] "The taking of Pskov"

## [38] "The Tale of Dracula"

## [39] "The tale of Igor's campaign"

## [40] "The Tale of Luka Kolocskij"

## [41] "The tale of the fall of Constantinople"

## [42] "Treaty of Alexander Nevskij and Novgorod with the Germans"

## [43] "Uspenskiij sbornik"

## [44] "Varlaam's donation charter to the Xutyn monastery"

## [45] "Vesti-Kuranty"

## [46] "Zadonshchina"

```

### Lemma:

```

## [1] "челов<U+0463>къ" "авъгустъ" "привьна" "архиепископъ" "гла
голъ"

## [6] "богородица" "апостолъ" "господинъ" "господь" "госуда
рь"

## [11] "м<U+0463>сяць" "мученикъ" "нед<U+0463>ля" "святитель" "сьрд
ьце"

## [16] "хрьстолюбьць" "ц<U+0463>сарица" "ц<U+0463>сарь" "ц<U+0463>сарьстви
е" "ц<U+0463>сарьство"

## [21] "епископъ"

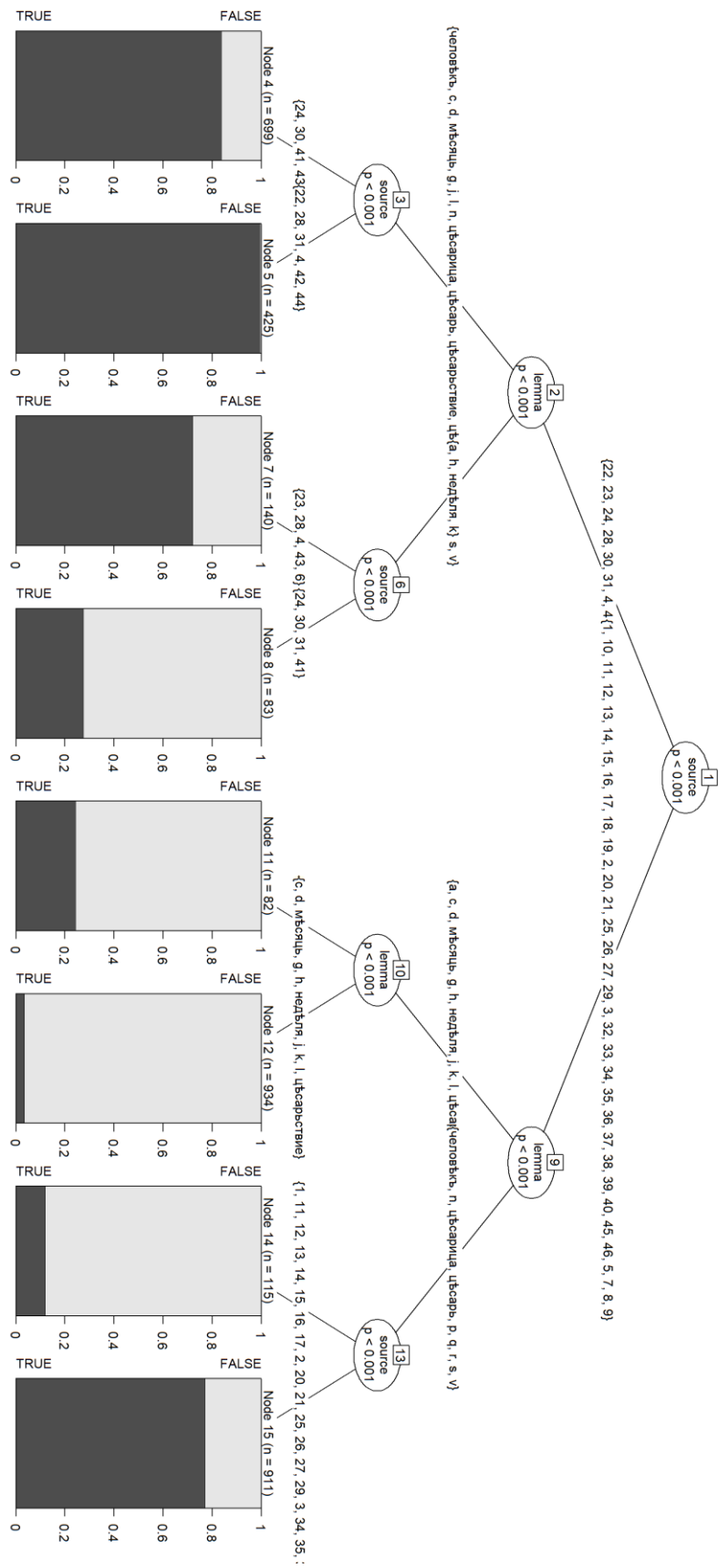
```

## 8.2 CART-Analysis & Random Forest

A CART-analysis seeks to make the cleanest split of the observations into several nodes based on the importance and interplay of the various variables. In this manner the CART-analysis determines the relative importance of the variables (Skjølsvold, 2021).

As the factors get printed above one another, we must consult the prose to make any sense of the splits in the CART-analysis, but in any case, we can observe the p value  $< 0.001$  for all (8) terminal nodes (Skjølsvold, 2021):

Figure 11 - CART analysis



(Skjølsvold, 2021)

```

##
## Conditional inference tree with 8 terminal nodes
##
## Response: has_titulo
## Inputs: source, lemma, case, number, length_difference, form_length
## Number of observations: 3389
##
## 1) source == {22, 23, 24, 28, 30, 31, 4, 41, 42, 43, 44, 6}; criterion =
  1, statistic = 1029.313
## 2) lemma == {челов<U+0463>къ, с, d, м<U+0463>сяць, г, j, l, n, ц<U+046
3>сарица, ц<U+0463>сарь, ц<U+0463>сарьствие, ц<U+0463>сарьство, p, q, r, s,
  v}; criterion = 1, statistic = 233.725
## 3) source == {24, 30, 41, 43, 6}; criterion = 1, statistic = 78.334
## 4)* weights = 699
## 3) source == {22, 28, 31, 4, 42, 44}
## 5)* weights = 425
## 2) lemma == {a, h, нед<U+0463>ля, k}
## 6) source == {23, 28, 4, 43, 6}; criterion = 1, statistic = 63.176
## 7)* weights = 140
## 6) source == {24, 30, 31, 41}
## 8)* weights = 83
## 1) source == {1, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 2, 20, 21, 25,
26, 27, 29, 3, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 46, 5, 7, 8, 9}
## 9) lemma == {a, с, d, м<U+0463>сяць, г, h, нед<U+0463>ля, j, k, l, ц<U
+0463>сарьствие, ц<U+0463>сарьство}; criterion = 1, statistic = 1035.682
## 10) lemma == {a, ц<U+0463>сарьство}; criterion = 1, statistic = 93.8
62
## 11)* weights = 82
## 10) lemma == {с, d, м<U+0463>сяць, г, h, нед<U+0463>ля, j, k, l, ц<U
+0463>сарьствие}
## 12)* weights = 934
## 9) lemma == {челов<U+0463>къ, n, ц<U+0463>сарица, ц<U+0463>сарь, p, q,
r, s, v}
## 13) source == {32, 33, 37, 38, 40, 45, 46}; criterion = 1, statistic
= 331.531
## 14)* weights = 115
## 13) source == {1, 11, 12, 13, 14, 15, 16, 17, 2, 20, 21, 25, 26, 27,
29, 3, 34, 35, 39, 5, 7, 8}
## 15)* weights = 911

```

##		locations							
##		4	5	7	8	11	12	14	15
##	FALSE	113	2	39	60	62	901	101	207
##	TRUE	586	423	101	23	20	33	14	704

As we can read from the prose related to the CART-Analysis, all splits into nodes dichotomizes between source and titlo, with a p value < 0.001: A regression model based on “has\_titlo ~ source + lemma” should be adequately equipped to predict whether a noun is abbreviated with a titlo or not. This result should be verified by a Random Forest (Skjølvold, 2021):

Figure 12 - Random Forest



(Skjølvold, 2021)

As it is evident from the graphic above, the result from the CART-analysis is indeed verified by the Random Forest, which also measures the importance of the variables. The Random Forest, like the CART-analysis also proposes the model “has\_titlo ~ source + lemma”, while discriminating against the other variables (Skjølvold, 2021).

**8.2.1 Prediction accuracy CART-Analysis & Random Forest**

When the tabulated predictions from the Random Forest is controlled for the actual values in the data frame, we get the following data (Skjølvold, 2021):

```

## $`1`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.8710747      0.1289253
##
## $`2`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.8710747      0.1289253
##
## $`3`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.8710747      0.1289253
##
## $`4`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.9671954      0.03280459
##
## $`5`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.9671954      0.03280459
##
## $`6`
##      has_titlo.FALSE has_titlo.TRUE
## [1,]      0.9845019      0.01549813

```

```

##           C           Dxy           n           Missing
## 0.9866807 0.9733614 3389.000000 0.0000000

```

By calculating the sum of the diagonal(s) divided by the total, we get prediction accuracy of the proposed model (Skjølsvold, 2021):

```

##
##           FALSE TRUE
## FALSE 1385 100
## TRUE   88 1816

```

The model which is suggested by the CART analysis and verified by the Random Forest make the right prediction (of whether a word is titolo-abbreviated or not) roughly 94% of the time (Skjølsvold, 2021).

```
## [1] 0.9445264
```

### 8.3 Generalized Logarithmic Regression Model(s), ANOVA & AIC

Fitting a Generalized Logarithmic Regression Model, which can predict when titolo abbreviations occur, is done by building a series of models and using ANOVA to measure which model performs best, while also observing the Akaike Information Criterion (AIC) for each of these models, to identify which is the better model (with the lowest AIC score) (Skjølsvold, 2021).

```
## Analysis of Deviance Table
##
## Model 1: has_titlo ~ 1
## Model 2: has_titlo ~ source
## Model 3: has_titlo ~ source + lemma
## Model 4: has_titlo ~ source + lemma + case
## Model 5: has_titlo ~ source + lemma + case + number
## Model 6: has_titlo ~ source + lemma + case + number + length_difference
## Model 7: has_titlo ~ source + lemma + case + number + length_difference +
##      form_length
##      Resid. Df Resid. Dev Df Deviance Pr(>Chi)
## 1      3388      4646.2
## 2      3343      3387.0 45  1259.19 < 2.2e-16 ***
## 3      3323      2003.1 20  1383.94 < 2.2e-16 ***
## 4      3315      1914.5  8    88.62 8.856e-16 ***
## 5      3313      1909.0  2     5.48 0.06451 .
## 6      3312      1785.4  1   123.61 < 2.2e-16 ***
## 7      3312      1785.4  0     0.00
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

AIC(jens.glm1)
## [1] 4648.216
AIC(jens.glm2)
## [1] 3479.03
AIC(jens.glm3)
## [1] 2135.095
AIC(jens.glm4)
## [1] 2062.474
AIC(jens.glm5)
## [1] 2060.992
AIC(jens.glm6)
## [1] 1939.38
AIC(jens.glm7)
## [1] 1939.38

```

Models 5, 6 and 7 have the lowest AIC scores, however, models 5 and 7 show  $p > 0.05$  or N/A. The summaries for models 5, 6 and 7 all have unidentified coefficients due to singularities. Model 4 however does not have any unidentified coefficients, but it fails the test for multicollinearity (Skjølsvold, 2021).

Model 3 “glm(formula = has\_titlo ~ source + lemma)”, which utilizes the same variables as proposed by the CART Analysis and Random Forest passes the test for multicollinearity, and as such the Generalized Logarithmic Regression Model is fitted. At the intercept we observe “has\_titlo (false) ~ source (Afanasij Nikitin’s journey beyond three seas) + lemma (человѣкъ)” with a p-value of  $< 0.05$  (Skjølsvold, 2021):

```

##
## Call:
## glm(formula = has_titlo ~ source + lemma, family = "binomial",
##      data = jens)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9620  -0.2784   0.1017   0.3954   2.7198
##

```



```

## Coefficients:
##
##          Estimate Std. Error z value Pr(>|z|)
## (Intercept)      3.86725    0.40595   9.526 < 2e-16 ***
## source10      -17.79756  6522.63862  -0.003 0.997823
## source11      -20.14405  2018.52229  -0.010 0.992038
## source12      -20.80684  2874.31129  -0.007 0.994224
## source13      -19.42230  1341.06269  -0.014 0.988445
## source14      -21.07114  2310.44055  -0.009 0.992723
## source15      -20.96369  2202.56489  -0.010 0.992406
## source16        -0.08409    0.39668  -0.212 0.832124
## source17         0.37814    0.40727   0.928 0.353155
## source18      -17.79756  4612.20201  -0.004 0.996921
## source19      -18.60468  6522.63862  -0.003 0.997724
## source2        -0.34596    0.67309  -0.514 0.607257
## source20       -3.86725    1.47133  -2.628 0.008578 **
## source21       -0.02427    1.48936  -0.016 0.986999
## source22       17.10966  2265.33373   0.008 0.993974
## source23       19.80347  4612.20202   0.004 0.996574
## source24        1.03988    0.50552   2.057 0.039681 *
## source25       -3.21081    0.87403  -3.674 0.000239 ***
## source26       -4.14458    1.09216  -3.795 0.000148 ***
## source27       -2.20160    0.71391  -3.084 0.002043 **
## source28        5.22383    0.47577  10.980 < 2e-16 ***
## source29        1.37913    1.29398   1.066 0.286509
## source3        -1.03122    0.78507  -1.314 0.188998
## source30        1.06997    0.39363   2.718 0.006564 **
## source31        1.81443    1.25217   1.449 0.147330
## source32       -4.16882    0.57606  -7.237 4.59e-13 ***
## source33       -4.27148    0.57032  -7.490 6.91e-14 ***
## source34       -1.67690    0.35948  -4.665 3.09e-06 ***
## source35       -2.09099    0.41869  -4.994 5.91e-07 ***
## source36      -17.27207  1452.79969  -0.012 0.990514
## source37      -18.94878   826.02559  -0.023 0.981698
## source38      -20.21189   776.55919  -0.026 0.979235
## source39      -20.87151  1490.05099  -0.014 0.988824
## source4        20.32240  4601.47393   0.004 0.996476
## source40      -22.32587  1385.95412  -0.016 0.987148

```

```

## source41          0.06422    0.42025    0.153 0.878537
## source42          18.52745 6522.63863    0.003 0.997734
## source43           0.99236    0.38714    2.563 0.010367 *
## source44          14.69882 6522.63861    0.002 0.998202
## source45          -20.79300 1378.53568   -0.015 0.987966
## source46          -21.92964 1251.30585   -0.018 0.986017
## source5           -1.20529    0.95825   -1.258 0.208462
## source6            1.30303    1.09911    1.186 0.235808
## source7          -20.05111 1911.82770   -0.010 0.991632
## source8           -1.03930    0.41720   -2.491 0.012733 *
## source9          -17.74727 1574.86343   -0.011 0.991009
## lemmaa            -5.10465    0.44172  -11.556 < 2e-16 ***
## lemmac            -3.82863    0.42992   -8.905 < 2e-16 ***
## lemmad            -4.63575    0.37762  -12.276 < 2e-16 ***
## lemmaм<U+0463>сяць -5.42179    0.34842  -15.561 < 2e-16 ***
## lemmaг            -6.23398    0.46978  -13.270 < 2e-16 ***
## lemmah            -5.07289    0.32203  -15.753 < 2e-16 ***
## lemmaнед<U+0463>ля -5.79928    0.40514  -14.314 < 2e-16 ***
## lemmaј            -6.12608    1.00452   -6.099 1.07e-09 ***
## lemmaк            -5.86403    0.41695  -14.064 < 2e-16 ***
## lemmaл            -5.76831    0.53888  -10.704 < 2e-16 ***
## lemmaп            -0.68778    1.06808   -0.644 0.519617
## lemmaц<U+0463>сарица -0.93318    0.68046   -1.371 0.170252
## lemmaц<U+0463>сарь -1.24261    0.27479   -4.522 6.12e-06 ***
## lemmaц<U+0463>сарьствие -3.68857    0.55782   -6.613 3.78e-11 ***
## lemmaц<U+0463>сарьство -4.16561    0.43260   -9.629 < 2e-16 ***
## lemmaр            1.08722    0.49959    2.176 0.029540 *
## lemmaq            -2.90919    0.26152  -11.124 < 2e-16 ***
## lemmar            -1.53330    0.50610   -3.030 0.002448 **
## lemmas            15.38355 1479.57124    0.010 0.991704
## lemmav            -2.35009    0.80000   -2.938 0.003308 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 4646.2  on 3388  degrees of freedom

```

```

## Residual deviance: 2003.1 on 3323 degrees of freedom
## AIC: 2135.1
##
## Number of Fisher Scoring iterations: 17

```

##	source10	source11	source12
##	1.000000	1.000000	1.000000
##	source13	source14	source15
##	1.000000	1.000000	1.000000
##	source16	source17	source18
##	3.769443	3.269957	1.000000
##	source19	source2	source20
##	1.000000	1.828958	1.080649
##	source21	source22	source23
##	1.122989	1.000000	1.000000
##	source24	source25	source26
##	5.862194	1.202572	1.145527
##	source27	source28	source29
##	1.450844	2.117564	1.107996
##	source3	source30	source31
##	1.235133	3.849852	1.082709
##	source32	source33	source34
##	1.648769	1.670430	7.123594
##	source35	source36	source37
##	2.972550	1.000000	1.000000
##	source38	source39	source4
##	1.000000	1.000000	1.000000
##	source40	source41	source42
##	1.000000	2.776363	1.000000
##	source43	source44	source45
##	4.634669	1.000000	1.000000
##	source46	source5	source6
##	1.000000	1.149498	1.089197
##	source7	source8	source9
##	1.000000	3.528690	1.000000
##	lemmaa	lemmac	lemmad

##	2.356086	4.235580	2.105610
##	леммам<U+0463>сяць	lemmaq	lemmah
##	2.204743	1.457253	2.561358
##	лемманед<U+0463>ля	lemmaj	lemmak
##	1.708360	1.099009	1.604915
##	lemmal	lemman	леммац<U+0463>сарица
##	1.505371	1.060622	1.122528
##	леммац<U+0463>сарь	леммац<U+0463>сарьствие	леммац<U+0463>сарьство
##	2.688995	1.258201	1.397771
##	lemmap	lemmaq	lemmar
##	1.376989	4.215246	1.289315
##	lemmas	lemmav	
##	1.000000		

As it is evident by the summary, the lemma is almost always significant (p value < 0.05 often), but source to a lesser extent. However, for some sources, it is highly significant in the prediction of the outcome (Skjølsvold, 2021).

### 8.3.1 Prediction accuracy Generalized Logarithmic Regression Model

When the predictions of the fitted Generalized Logarithmic Regression Model is controlled for the actual values in the data frame, we get the following table, where the sum of the diagonal(s) divided by the total yields the prediction accuracy (Skjølsvold, 2021):

##		FALSE	TRUE
##	FALSE	1316	169
##	TRUE	185	1719

The generalized logarithmic regression model predicts the right outcome in roughly 89% of the examples (Skjølsvold, 2021).

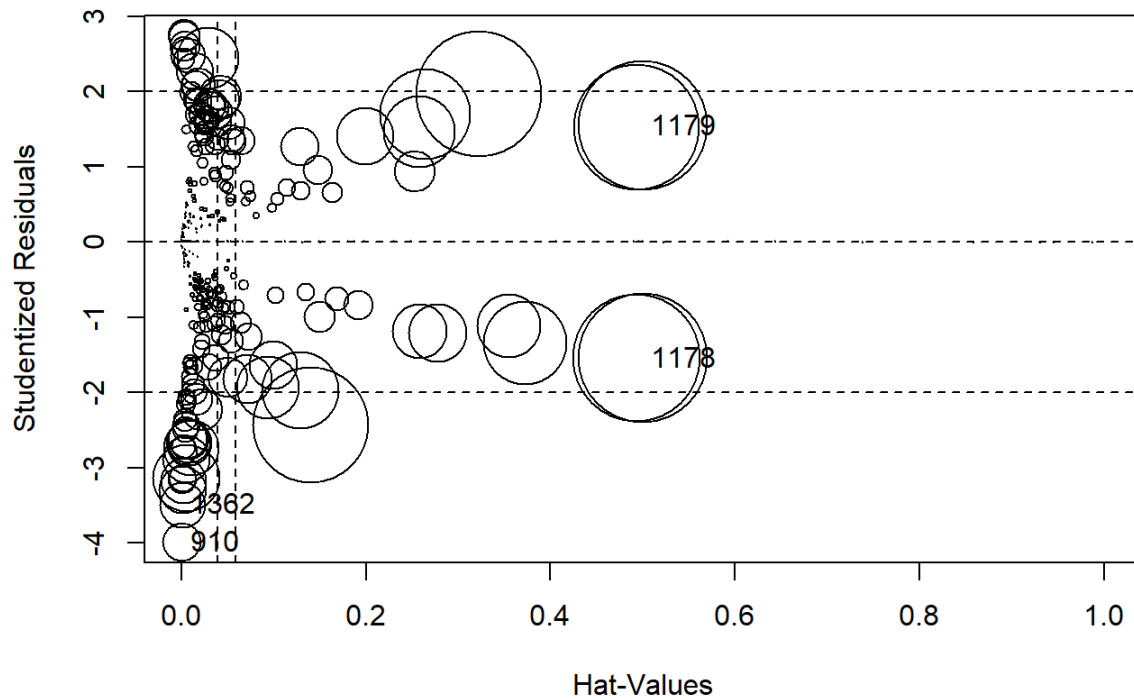
```
## [1] 0.8955444
```

The GLM model has lower prediction accuracy (roughly 89%) than the model predicted by CART (roughly 94%). However, this is a relatively small difference, and the GLM makes for a good model to predict whether a noun is titlo-abbreviated or not (Skjølsvold, 2021).

## 8.4 Outliers

An Influence Plot identified the following outliers:

Figure 13 - Influence Plot



(Skjølsvold, 2021)

##	StudRes	Hat	CookD
## 910	-3.984485	6.962945e-05	0.002703081
## 1163	NaN	1.000000e+00	NaN
## 1178	-1.544764	5.000000e-01	0.030303030
## 1179	1.544764	5.000000e-01	0.030303030
## 1362	-3.485454	6.517946e-04	0.003782771
## 1660	NaN	1.000000e+00	NaN

Of these, observations 910 and 1362 have greater or smaller studentized residuals than 2 and -2, respectively, however, their HAT-values are low, along with Cook's Distance. An inspection of these two observations reveal that they account for lemmas that were expected to be titlo-abbreviated but were not, цѣсарь (tsar) and богородица (mother of God), respectively (Skjølsvold, 2021).

## 8.5 Interpretation of the statistical analysis

Arguably, it might seem intuitive that either the length of a word, or how much a word could be shortened, would motivate titlo-abbreviations. It is however evident by the statistical analysis. Neither the CART analysis, Random Forest or General Logarithmic Regression models supports this. Whether a noun is abbreviated by titlo or not is dependent on source and lemma, that is, whether a noun was titlo-abbreviated or not was dependent on the author and what word the author was writing. The evidence from the statistical analysis upholds what is described in the literature, that titlo-abbreviations were used to convey meaning – they were not used as a pragmatic solution in face of scarcity of writing materials, or for the purpose of efficiency (Skjølsvold, 2021).

## 9 Discussion

As we have learned from the literature review, writing materials were both expensive and limited (c.f. Kaldor), which may have been an incentive to utilize titlo-abbreviations. Also, the transition from *ustav* to *poluustav* (and *skorpis*’) allegedly promoted the use of titlo-abbreviations in Muscovy (c.f. Tikhomirov), as Old East Slavic scribes also had an affinity for the use of titlo-abbreviations (c.f. Matthews). However, the result from the statistical analysis lends no support for the alternative hypothesis (c.f. section 5), that simple titlos were utilized due to an economical incentive, or for other pragmatic reasons, since the factors *length difference* and *form length* are insignificant, and not a part of the best fitted regression model (c.f. section 6.2.1 for an overview of all the factors).

We should keep in mind, that abbreviations without titlos were also utilized, and that that dataset used in in the statistical analysis is composed of word forms shorter than the lemma, that is, abbreviations without titlos at all - these may well have been pragmatic, but abbreviation by titlo was reserved to convey a special (sacred) meaning (c.f. Zhivov & Uspenskij). We have learned from the statistical analysis, that the factors *source* and *author* allow us to predict whether a noun would be titlo-abbreviated or not. The importance of the factor *author* can also be linked theorizations of semiology (c.f. Saussure & Harris), in terms of *signans* and *signatum* (I will however refrain from explicating upon this). Furthermore, many titlo-abbreviations simply involve shortening the word by a single letter. We are left without any convincing arguments or evidence in favor for economical or pragmatic use of titlo-abbreviations. When it comes to titlos – it’s all about meaning.

As we have learned from the theory section, abbreviations themselves are metonymical, c.f. Brdar. Thus, synthesizing Cognitive Linguistics theory with the interpretation of titlos by Zhivov & Uspenskij, it is evident that titlo-abbreviations index a sacred meaning.

If we consult Gammanovich's list of frequently titlo-abbreviated words in (Old) Church Slavonic (c.f. figure 2 - Gammanovich's list of common titlo-abbreviations), there is some overlap with Old East Slavic, as evident by the data used in this thesis. E.g., the (Old) Church Slavonic equivalents of апостоль, господь, епископъ, мученикъ, недѣля, святитель, цѣсарьство, and цѣсарь are all listed by Gammanovich, which attest to their sacred attributes (when titlo-abbreviated). However, some of the most common titlo-abbreviations in Old East Slavic, e.g., гривна, августъ, and мѣсяць do not bring about any obvious sacred meaning at first glance. I suggest that we may infer their special meaning, or their contiguity to the prototype, and classify them typologically (c.f. Nessel, 2015), by looking closer at selected observations in context, and by juxtaposing them.

The discussion herein relies heavily on introspection (c.f. Janda, 2013), and it is highly speculative, but I will attempt to infer how all the most frequently titlo-abbreviated nouns relate to the prototypical meaning of sacred through contiguity and categorize them according to typologies in the following sections. The latter will allow me to draw up a radial network, which visualizes their contiguity to the prototypical meaning of sacred.

## 9.1 Meaning

Having determined what titlo-abbreviations are not - that is, a pragmatic diacritic, but instead a diacritic used to index sacred meaning, we may be able infer this meaning by looking at titlo-abbreviations in context.

Some nouns, such as гривна, which hardly seem sacred, may have been subject to such a high frequency in (secular) texts, that it became normative to abbreviate it, e.g.:

(170133) а за зубъ *зрѣ*

and grivna(s) for a tooth.

This observation speaks about the debt owed for injury. Now, it seems highly unlikely that a monetary unit in such a context would have any sacred meaning about it. However, this observation is drawn from *Russkaya Pravda*, the legal code of Rus' and its principalities,

which were ruled by monarchs who were juxtaposed to God (c.f. Uspenskij & Zhivov). I argue that we may therefore infer contiguity from *цѣсарѣство* and *цѣсарѣствіе*:

(125037) *ныне же увѣдѣлъ есть князь нашъ послати грамоту ко **црѣву** нашему*

now our prince found out to send a document to our kingdom.

In this example the realm of the monarch is regarded as sacred, evident by the titlo-abbreviation, and thus we have a rationale, albeit a weak one, in support of the sacred nature of *гривѣна* through contiguity. However, if we look closer at *цѣсарѣствіе* the metonymic extension from the prototype becomes even more clear:

(172903) *его же **црѣствию** нѣсть конца*

His kingdom has no end

The titlo-abbreviation indexes God's Kingdom, which is undoubtedly sacred (in the right context), while the monarch's kingdom is sacred through a metonymic extension from God's kingdom.

If we look at typologies (c.f. Nessel, 2015), *цѣсарѣствіе* and *цѣсарѣство*, arguably, relate to containment, while *гривѣна*, belongs to (a new) category derived from Containment through contiguity, a category I refer to as *law* (due to the origin of *гривѣна*, in *Russkaya Pravda*).

While it is speculative, we can infer how *гривѣна* can be viewed as sacred by looking closer at its context and its contiguity, or through metonymic extensions. Yet, this only explains one of the problematic observations, but we may infer the meaning of the remaining troublesome observations, by using the same logic and looking into contiguity relations.

While we do find *недѣля* in Gammanovich's list of words often abbreviated by titlo in (Old) Church Slavonic, we must keep in mind this thesis concerns itself with Old East Slavic, as such there may be discrepancies. Looking into an example of *недѣля* in context, we observe that it concerns itself with a cyclical event in the liturgical calendar:

(212591) *и въниде на сборъ по чстѣи **недѣ***

and entered (into) the cathedral in the first week of Lent.



Here, it makes perfect sense that недѣля is abbreviated by titlo to index its special, sacred meaning as the first of week Lent. But does this apply to the other cyclical events, or nouns related to time in general?

(133490) *бѣ бо тогда мѣць грудень рекше ноябрь*

for then it was the month of Gruden, which is called November.

When we look to мѣсяць, the contiguity is not so obvious. In this observation, we observe that мѣсяць is titlo-abbreviated. However, in terms of chronicles and annals, мѣсяць may have been subject to abbreviation by its sheer frequency, and thus norm. Moreover, months are cyclical events that occur repeatedly (like Lent), but this is hardly a satisfactory explanation, and it is hard to infer a special meaning. However, looking into the next observation, we see that both мѣсяць and авѣгустъ is titlo-abbreviated:

(214561) *и прїде в новѣгородъ мѣца авѣгусѣвъ а*

and (s)he came to Novgorod on the first day of August.

Note, how neither the proper name Gruden, nor November was titlo-abbreviated in the earlier example. If look back to the etymology of tsar, we know that it is derived of Caesar. It is perhaps not common knowledge, but widely known, that the month of August was named after Caesar Augustus. Now, it is uncustomary or even bad conduct to introduce new citations in a discussion, so I do apologize, but if we look to the etymology of August (or Augustus), one of its meanings was, or is, *venerable* (Augustus, 2023). This might explain why the month of August, in the proper context, could be titlo-abbreviated, given its origin. If the month was named after, or in honor of Caesar, it may be considered sacred through contiguity, in the right context. This reasoning, however, is somewhat anachronic. It is not certain that the Old East Slavic scribes in medieval Rus were aware of the etymology of August, save for a select few. Therefore, it is most speculative.

We're left without a satisfactory sacred meaning of мѣсяць or авѣгустъ, but it is however not improbable that it may still be considered sacred in some contexts. Though, as already mentioned, some nouns may have occurred at such a high frequency, that abbreviating these nouns by titlo became normative. This may apply to мѣсяць and авѣгустъ in terms of chronicles and annals and, perhaps, also due to contiguity to numbers (c.f. Gammanovich).

August may simply have been an eventful month through the course of years, decades, and even centuries. Alas, there are exceptions to every rule.

However, we may categorize these nouns (недѣля, мѣсяць, and августъ) typologically in a category based upon cyclical time, or the liturgical calendar, derived from the typology of adjacency. I will return the rationale of this categorization further below.

We're still left with other somewhat ambiguous observations, that is сръдце and глаголь, but they become quite clear once they are put into context and juxtaposed with other observations.

First:

(125607) *и приложиши въ сръце твое в разумъ*

and you accept in your heart wisdom.

Here, we observe that virtue is stored in the heart, and further we observe, that man also is titlo-abbreviated:

(160544) *яко члѣвкъ оумираеть*

that man dies

The sacred nature of man (when titlo-abbreviated), as prescribed by Sazonova (c.f. section 2.1), should also apply to man's heart, through contiguity. And I argue, that сръдце and члѣвкъ should be categorized as part-whole relations to the prototype, c.f. section 2.1.

Though, сръдце could also be viewed as a case of containment, but I infer contiguity from the prototype through члѣвкъ. Now we're left with the final ambiguous titlo-abbreviation, глаголь:

(192271) *иди по глѣбу моему*

go with my word.

Which can be juxtaposed with the following two observations, to elucidate that is refers to God's word:

(126004) *и реч ѡръ*

and the Lord said

(126963) *и речъ г҃ъ*

and the Lord said

The Lord, who's indexed meaning is God when titlo-abbreviated, is sacred, and just like he sacred, his words may also sacred, which explains why also глаголь would be titlo-abbreviated in the right context. This also further alludes the perpetuity of God's word. I argue that глаголь should be categorized typologically as adjacency, whereas цѣсарь and господь should be categorized as part-whole relations in terms of the relation to the prototype, as the nouns refer to God (who's indubitably sacred) in this context.

Now, I may explicate upon how недѣля, мѣсяць, and августъ have been categorized, by looking to глаголь, which is typologically categorized as adjacency. God's word, which is sacred, decrees how (and when) man should worship him, which justifies deriving a category of cyclical time, or the liturgical calendar, from adjacency, where we find глаголь.

Moreover, other observations that speak of God also testify to titlo-abbreviations:

(159574) *милыи г҃нѣ наю и драгыи*

Our merciful and graceful God!

(185386) *батюшко г҃дрь*

O father sovereign!

These examples, господинъ and государь along with цѣсарь and господь, all relate to a part-whole relationship with the prototype, as all these words indices the meaning God, when titlo-abbreviated.

Now we depart into two (more) fascinating cases of titlo-abbreviations, in the form of цѣсарица and богородица:

(127545) *и посла к нему цѣрица рькуще*

and sent to him the tsarina saying

(157699) *и соблюди ѿ всако плѣненья вражъя твои градъ бѣѣ*

and save your city from all hostile occupation, O Mother of God!

In terms of typology, we should bear in mind that цѣсарица and богородица were the vessels that gave birth to God (as Jesus) and Tsars, respectively, and are thus examples of containment (much like цѣсарѣствие & цѣсарѣство, however, not the same). The latter, богородица, also invoke the purported Muscovite use of titlos in terms of homonymy and polysemy with regards to proper names (c.f. Collins in section 4.2), which also would apply to the names of apostles, martyrs, and saints:

(186999) *и слушаи что пророкъ говорит со **апѣлом***

and listening to what the prophet says to the apostle

It is curious how пророкъ (186999) is not titlo-abbreviated in terms of the context. Perhaps we are dealing with a prophet, but not *the* prophet, but it cannot be deciphered in this context, but I digress.

(208079) *призва на помочь собѣ стаг **мчѣика** фѣвдора*

and called the holy martyr Fyodor to help

We should note that the apostle is explicitly stated to be holy.

(198795) *или самъ боуди игоумень или шед сѣпроси намъ игѣмена оу **стѣла***

Either be the abbot yourself or go and ask the saint to appoint an abbot for us

It is curious how abbot is not titlo-abbreviated, but, not surprising that апостолѣ, мученикѣ, and святителѣ are titlo-abbreviated as they are arguably of high que validity in terms of sacred(ness). I argue that they should be categorized typologically in terms of contact, due to their communion with God. We may also draw parallels to their roles as disciples and their passion, but such arguments may diverge into other domains, and thus be a case of metaphor, rather than metonymy.

In terms of the category of cyclical events, we should also note, that the apostles, martyrs, and saints are (sacred) objects that are worshipped during their respective feasts, according to the liturgical calendar.

(214061) *и приде поставленъ **архѣпсѣ** антонию*

and there came the appointed Archbishop Anthony

(205865) *и не да епѣтъ надъ нима пѣти*

and may not the bishop sing (praise) over them

Considering the role of the clergy (and their relationship to God), it's hardly surprising that they are titlo-abbreviated, like the apostles, martyrs, and saints. In terms of their special sacred meaning, it probably revolves around their subjugation to (the true) God, and not some idol (c.f. Gammanovich), but as in the case above (c.f. апостоль, святитель, and мученикъ), we're now threading the waters of metaphor, and I will refrain from digressing.

Typologically, they may also, like апостоль, святитель, and мученикъ, be categorized in terms of Contact, through their communion with God.

The following observation could also be tied to the sacred meaning of человекъ and сердце, but I have likened христолюбыцъ to apostles, martyrs, and saints with regards to typology, due his (or her) communion with God. In terms of the indexed sacred meaning, though, it is hard to see that a lack of titlo would change much:

(192284) *таче пакы дъждю прѣставъшию ѿиде хлюбъцъ въ домъ свои*

when the rain stopped the Lover of Christ went to his home

Arguably, we're able to infer contiguity from the prototypical meaning of the titlo, that is, sacred, and apply it to all the most frequently titlo-abbreviated nouns. Though, some of the contiguity relations or metonymic extensions drawn above are highly speculative, such as the case of мѣсяць and августъ.

## 9.2 Radial Network

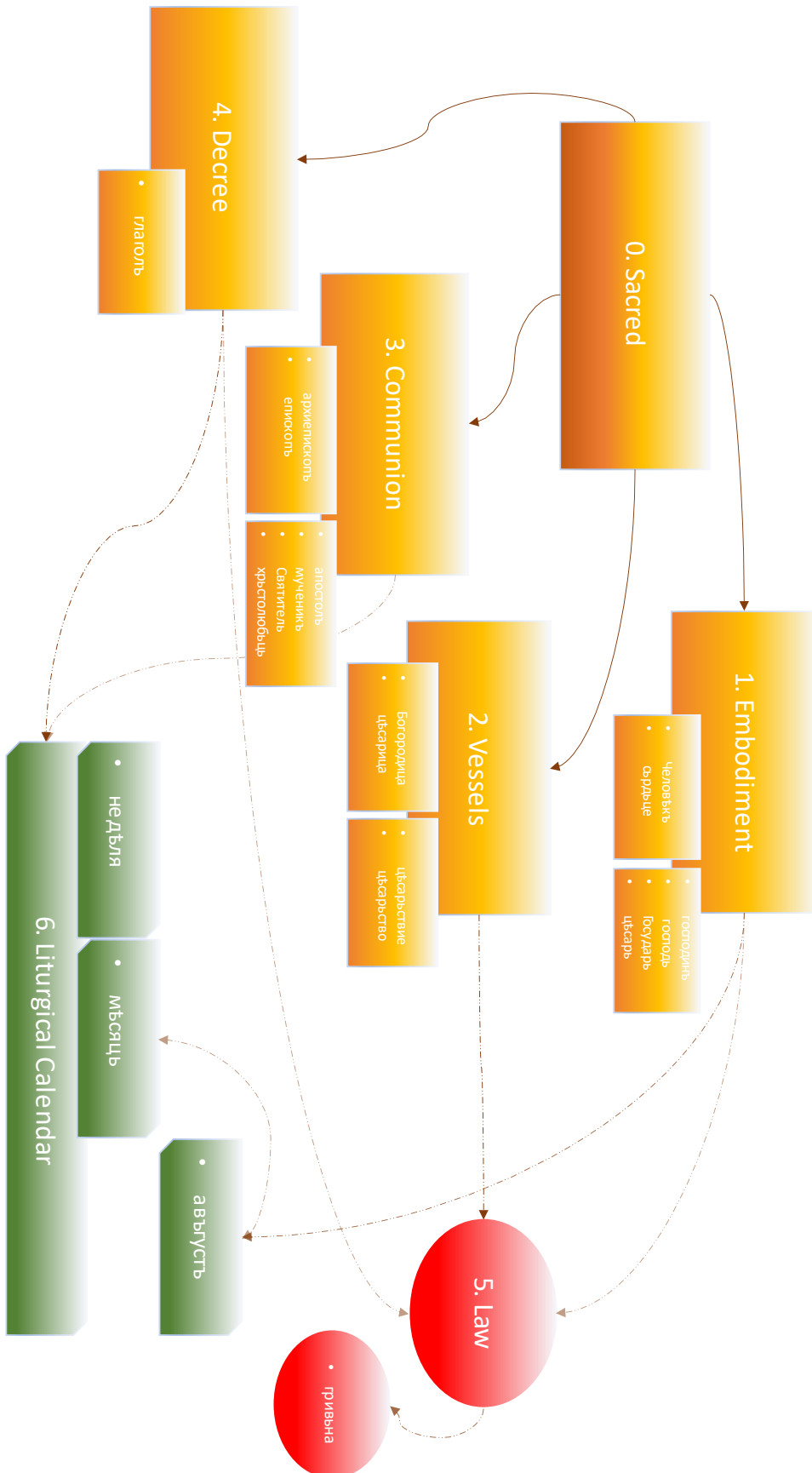
I propose a radial network with categories drawn upon contiguity from the prototype - sacred, inspired by Peirsman and Geeraerts' Typology of Metonymy as presented by Nessel, shown in figure 13 – Radial network below.

The radial network, centered around the prototype (sacred), diverges into the metonymically typological categories of 1) Embodiment (part-whole), 2) Vessels (containment), 3) Communion (contact), and 4) Decree (adjacency). Contiguity, or metonymic extension, is demonstrated by lines between the categories. I will refrain from reiterating the typological categorization of the titlo-abbreviated nouns in this section, as they are superimposed upon the radial network.

Arguably, we might argue that the nouns are also categorized by the magnitude of how sacred they are, c.f. figure 9 - Typology of metonymical relations.

The discussion preceding this radial network is, as already stated, speculative, and as such this radial network is only a suggestion – further, targeted, research is necessary to make any solid statements about contiguity relations or metonymic extensions derived from the prototypical meaning and its (radial) networks. This is however a starting point.

Figure 14 - Radial Network



### 9.3 Further research

Further quantitative research into titlo-abbreviations should, if possible, control for the distribution of titlo-abbreviations according to time and their origin by region, if not also classify writings as either ecclesiastic or secular. Arguably, it could be interesting to research the relative use of titlo-abbreviations in different sources, as the sheer volume of selected tomes may obfuscate findings.

However, we find claims in the literature that both *simple* and *lettered* titlos were utilized due to economic incentives (c.f. Tikhomirov). As such, it may be purposeful to research *lettered* titlos specifically in a manner akin to this thesis.

## 10 Conclusion

Considering the results of the statistical analysis and discussion above, we may readily conclude what titlo-abbreviations are not, that is – a pragmatic tool. There's no evidence in support of titlo-abbreviations being utilized for pragmatic purposes in terms of economizing writing, with regards to materials and penmanship. While we find arguments in favor of this view in the literature, it is evident by statistical analysis that the length of word abbreviated, and the length difference between the titlo-abbreviation and the full form did not motivate scribes to use titlo-abbreviations. Further scrutiny of several observations reveals that abbreviation by titlo oftentimes only meant shortening the word by a single letter.

Thus, we can state with confidence that titlo-abbreviations are all about meaning. The analysis in this thesis is not equipped to make absolute statements about the meaning of titlos, but by scrutinizing selected observations, we find support for the view presented by Zhivov & Uspenskij in *Tsar and God and Other Essays in Russian Cultural Semiotics*, that titlo-abbreviations were utilized to denote how an is object sacred, which can be described as the prototypical meaning of titlos in the scope of Cognitive Linguistics.

The paradigmatic approach utilized in this thesis, drawn from Cognitive Linguistics, has proved itself applicable to historical (or diachronic) linguistics research. However, future research into titlos specifically may benefit from some tuning or methodological adjustments compared to the statistical analysis carried out in this thesis, with regards to the data frame.

To wrap this thesis up, I would like to reiterate a part of a statement made by Nessel's in *When We Went Digital and Seven Other Stories about Slavic Historical Linguistics in the 21st*



*Century* (2017, p. 439): (...) *historical work in Slavic is still alive and kicking in the 21<sup>st</sup> century (...).*

## Works cited

- Allaire, J. J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. (2012). *rmarkdown: Dynamic Documents for R. R package version 2.7*. <https://rmarkdown.rstudio.com>
- Augustus. (20.05.2023). In Wiktionary. <https://en.wiktionary.org/wiki/augustus#Latin>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Brdar, M. (2017) *Metonymy and Word-Formation: Their Interactions and Complementation*. Cambridge Scholars Publishing.
- COLLINS, D. (1992). ON DIGLOSSIA AND THE LINGUISTIC NORMS OF MEDIEVAL RUSSIAN WRITING. *Studies in Slavic and General Linguistics*, 17, 79-94. Retrieved from <http://www.jstor.org/stable/40997021>
- Dabrowska, E., & Divjak, D. (2015). *Handbook of Cognitive Linguistics: Vol. Band 39*. De Gruyter, Inc. <https://doi.org/10.1515/9783110292022>
- Eckhoff, H. M., & Berdicevskis, A. (2015). Linguistics vs. digital editions: The Tromsø Old Russian and OCS Treebank. *Scripta & e-Scripta 14–15*, pp. 9-25.
- Ferguson, C. A. (1959). Diglossia. *Word (Worcester)*, 15(2), 325–340. <https://doi.org/10.1080/00437956.1959.11659702>
- Fox, J., & Weisberg, S. (2019). *An {R} Companion to Applied Regression*, Third Edition. Thousand Oaks CA: Sage. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- Gammanovich, A. (2001). *Grammar of the church Slavonic language* (p. 438). Jordanville, NY: Holy Trinity Monastery.
- Ginsburgh, & Weber (2016). *The Palgrave Handbook of Economics and Language*. Palgrave Macmillan UK.
- Grannes, A. (1989). История русского литературного языка (XI-XVII вв [Review of *История русского литературного языка (XI-XVII вв.)* by B. A. Uspenskij]. *Russian Linguistics*, 13(3), 261–269. <http://www.jstor.org/stable/40160261>

- Harrell Jr., F. E. (2021a). *Hmisc: Harrell Miscellaneous*. R package version 4.5-0.  
<https://CRAN.R-project.org/package=Hmisc>
- Harrell Jr., F. E. (2021b). *rms: Regression Modeling Strategies*. R package version 6.2-0.  
<https://CRAN.R-project.org/package=rms>
- Haser, V. (2005). *Metaphor, Metonymy, and Experientialist Philosophy : Challenging Cognitive Semantics*, De Gruyter, Inc.
- Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006). Survival Ensembles. *Biostatistics*, 7(3), 355--373.
- Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3), 651--674.
- Ivanov, V. V. (2010). Dvoeverie. In *The Oxford Dictionary of the Middle Ages*: Oxford University Press. Retrieved 22 Feb. 2023, from <https://www-oxfordreference-com.mime.uit.no/view/10.1093/acref/9780198662624.001.0001/acref-9780198662624-e-1870>
- Janda, L. A. & Dickey, S. M. (2017). Cognitive Linguistics. *Journal of Slavic Linguistics*, 25(2), 367–386. <https://doi.org/10.1353/jsl.2017.0014>
- Janda, L. A. (2013). *Cognitive Linguistics: The Quantitative Turn. The Essential Reader*. Berlin: De Gruyter Mouton.
- Kaldor, I. L. (1967). *Slavic Paleography and Early Russian Printing: The Genesis Of The Russian Book*. ProQuest Dissertations Publishing . Retrieved from <https://www.proquest.com/dissertations-theses/slavic-paleography-early-russian-printing-genesis/docview/302285958/se-2?accountid=17260>
- Levshina, N. (2015) *How to do Linguistics with R. Data exploration and statistical analysis.*, John Benjamins Publishing Company. <http://ebookcentral.proquest.com/lib/tromsoub-ebooks/detail.action?docID=4386605>
- Lunt, H. G. (1990). History, Nationalism, and the Written Language of Early Rus'. *The Slavic and East European Journal*, 34(1), 1–29. <https://doi.org/10.2307/309309>

- Matthews, W. (1953). The Russian Language before 1700. *The Slavonic and East European Review*, 31(77), 364-387. Retrieved from <http://www.jstor.org/stable/4204458>
- Nesset, T. (2015). How Russian came to be the way it is: a student's guide to the history of the Russian language (pp. XXV, 361). Slavica.
- Nesset, T. (2016). “Cyclic” time in the history of Russian: Culture and language internal factors. *Studies in Language* 40.3, 591-621.
- Nesset, T. (2017). When We Went Digital and Seven Other Stories about Slavic Historical Linguistics in the 21st Century. *Journal of Slavic Linguistics*, 25(2), 439–462. <https://doi.org/10.1353/jsl.2017.0017>
- Nesset, T. (2022). *Language Change and Cognitive Linguistics: Case Studies from the History of Russian (Elements in Cognitive Linguistics)*. Cambridge: Cambridge University Press. doi:10.1017/9781009031554 <https://www-cambridge-org.mime.uit.no/core/elements/language-change-and-cognitive-linguistics/7D2DF80D0DD29714758DDE5BE81B4430>
- Sarkar, D. (2008) *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN 978-0-387-75968-5
- Saussure, F., & Harris, R. (2021). *Course in general linguistics (First ed., Bloomsbury Revelations)*. London, England: Zed Books. <http://ebookcentral.proquest.com/lib/tromsoub-ebooks/detail.action?docID=1507640>
- Sazonova, Н. И. (2006). ПРАВОСЛАВНЫЙ БОГОСЛУЖЕБНЫЙ ТЕКСТ И СОЦИАЛЬНЫЕ АСПЕКТЫ ФУНКЦИОНИРОВАНИЯ ПРАВОСЛАВНОГО ПРЕЛИГИОЗНОГО СОЗНАНИЯ in Вестник ТГПУ, 12 (63). Серия: ГУМАНИТАРНЫЕ НАУКИ (СОЦИОЛОГИЯ). <https://cyberleninka.ru/article/n/pravoslavnyy-bogosluzhebnyy-tekst-i-sotsialnye-aspekty-funksionirovaniya-pravoslavnogo-religioznogo-soznaniya/viewer>
- Skjølsvold, J. K. (2021). *Titlos in Old Russian literature [HIF-3082 Exam Paper, University of Tromsø]*.

- Strobl, C., Boulesteix, A., -L, Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional Variable Importance for Random Forests. *BMC Bioinformatics*, 9(307).
- Strobl, C., Boulesteix, A., -L., Zeileis, A., & Hothorn, T. (2007). Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*, 8(25).
- Ungerer, F., & Schmid, H.-J. (2006). *An Introduction to Cognitive Linguistics* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315835396>
- Uspenskij, B. (1987). *Istorija russkogo literaturnogo jazyka (XI-XVII vv.): (Az orosz irodalmi nyelv toerténete a XI-XVII. században)*. Na russ. jaz. Peter Lang International Academic Publishing Group.  
<https://directory.doabooks.org/handle/20.500.12854/27346>
- Wickham et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686, <https://doi.org/10.21105/joss.01686>
- Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wickham, H., Francois, R., Henry, L., & Muller, K. (2021). *dplyr: A Grammar of Data Manipulation*. R package version 1.0.5. <https://CRAN.R-project.org/package=dplyr>
- Xie, Y., Allaire, J. J., & Golemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC. ISBN 9781138359338. <https://bookdown.org/yihui/rmarkdown>.
- Xie, Y., Dervieux, C., & Riederer, E. (2020) *R Markdown Cookbook*. Chapman and Hall/CRC. ISBN 9780367563837. <https://bookdown.org/yihui/rmarkdown-cookbook>
- Zhivov, V. & Uspenskij, B. (2012). *Tsar and God and Other Essays in Russian Cultural Semiotics*. Academic Studies Press. <http://ebookcentral.proquest.com/lib/tromsoub-ebooks/detail.action?docID=3110499>
- Архивы ргада о народах Забайкалья и реки Шилка. (2022). *Siberian Research*, 7(1), 31–32. <https://doi.org/10.33384/26587270.2022.07.01.05r>

Козлов, М. Н. & Матвеева Я. В. (2021). Культ двуликих божеств в восточнославянском язычестве in *Voprosy istorii*, 12(2), p. 114-121

Тихомиров, М. Н., & Муравьев, А. В (1966) Развитие русского кирилловского письма.

Retrieved from:

<https://web.archive.org/web/20081201132619/http://www.opentextnn.ru/history/paleography/tihomirov/?id=1493>

Царь. (17.05.2023). In Wiktionary.

<https://en.wiktionary.org/wiki/%D1%86%D0%B0%D1%80%D1%8C#Russian>

