

Martin of Cochem's Golden Key of Heaven and its Czech Relatives: Quantitative Analysis of Baroque Prayers

Keywords: Martin Cochem, Manuscript, Náchod, quantitative analysis

Abstract

Rewriting books was a widespread phenomenon during the Baroque period of the Czech literature. The manuscripts were not always „honest copies”, on the contrary, scribes often compiled several sources or added their own texts to the original. The famous book *Golden Key of Heaven* by Martin of Cochem is compared with a manuscript *Key of Heaven* from a Regional museum in the town of Náchod. We use two statistical methods, both of them strongly indicate that the manuscript is a copy of some chapters from the *Golden Key of Heaven* rather than a compilation.

1. Introduction

The period of the Baroque literature in the Czech countries, which started around 1640s and lasted through 1760s (Lehár 1998, Vašica 1995), is traditionally perceived through the prism of two religious de-

nominations, namely Catholic and Protestant (mostly in exile). This violent division of Czech literature is a consequence of the defeat of the anti-Habsburgs opposition in the Battle of White Mountain (1620). Soon after the battle, a new constitutional law, the so-called Revised Provincial Land Ordinance (1627), allowed only the Catholic religion in the Czech countries, and established the hereditary right of the Habsburgs to the Czech throne.

The period of the intensive recatholization and the influential Baroque spirituality resulted in a huge increase of the Czech Catholic religious literature; genres like homilies, spiritual songs, hymn-books, guidebooks for pilgrims, and prayer books dominated. In fact, printed prayer books and later published manuscript prayer books were extremely popular and were issued until the 19th century.

We focus on manuscripts (as opposed to printed books) in this paper (and on one manuscript in particular). Since the Reformation, literacy has increased among ordinary people in Europe. This process was accompanied with a growing book trade. The trade had several manifestations, from printing and selling as we know it today to – from the present perspective uncommon – rewriting. Especially in rural areas, rewriting was a cheaper way to obtain a book until the 19th century for a majority people. Consequently, rewriting was a business, albeit small. As in any business, a seller needed to market products and a buyer wanted to buy good ones. Further, as in any business, there existed some „tricks” among sellers (like, e.g., titles were deliberately chosen to be similar to bestsellers, see below).

The *Golden Key of Heaven*² (translated from German original *Guldener Himmelschlüssel*) by Martin of Cochem is considered to be the most often published prayer book from the period. Its author, a German Capuchin who spent two years in a monastery in Prague, wrote his texts in an understandable, meditative language full of

² Its first translation to the Czech language appeared in 1701. It had been reprinted unchanged several times until 1775, when the texts were influenced by the censorship of the Imperial Office.

¹ J. Mačutek was supported by research grant VEGA 2/0096/21.

imagination, and at the same time without dialectal variations and assaults on Protestants.³ Cochem's prayers reflect not only liturgy, but also topics from the everyday life of the believers, especially women (birth, education of children, marriage etc.). His *Golden Key of Heaven* and books of folk reading influenced both the rural and city spirituality of the 18th and 19th century in a dimension which is today difficult to imagine (Martin 2007, p. 8).

The label *Key of Heaven* became a symbol of a „good content” in the discourse of the Czech early modern age; therefore, it was often abused by publishers, as it was a guarantee of a good sale. New versions⁴ of the book were published, with titles like, e.g., *Little Golden Key of Heaven*, *Half Key of Heaven* etc., in which Martin of Cochem is not mentioned as the author.

The same phenomenon can be observed in the manuscript production. Manuscripts were written on contract by churchmen, teachers, peasants etc., who re-used some texts from Cochem's prayer books. *Repertory of manuscripts of the 17th and 18th century from Czech museums' collections* (Linda 2003a, 2003b; Fidlerová, Bekešová 2007a, 2007b) contains hundreds of manuscript prayer books, with many of them referring to the production of Martin of Cochem. As it was believed that they were direct copies of printed books with which they shared the same titles, those manuscripts were considered to be only of a marginal scientific interest. However, more detailed analyses suggest that, in some cases, the printed books written by Martin of Cochem and their would-be copies are very similar in some aspects, but the manuscripts are most probably influenced also by another author or by the scribe's own invention. We present one pair of such texts as an example.

³ According to Jan Kvapil (2001, pp. 42–43) these were the reasons of its popularity.

⁴ We understand the term *version* in a textual sense as a variability in a thematic, content, compositional and language construction of publication (Flaišman – Kosák, 2006).

Table 1. Text in the printed book and

Printed book	Manuscript
Při skonání tohoto dne srdce mé k tobě pozdvihuji, Pane nebe i země, a za všechna dobrodiní, která jsi mi dnes prokázal a že jsi mne na duši i na těle ode všeho zlého chrániti ráčil, srdečně ti děkuji. [...] (<i>Zlatý nebeklíč</i> , Orlické muzeum Choceň, Sbs 40, 1716, p. 20) ⁵	Při konci toho dne pozdvihuji můj mysl k tobě, ó Bože nebe i země, a děkuji srdečně tobě za všechno dobrodiní, které jsi nám dnes z milosti prokázal a od zlého, jak tělesného, tak duchovního otcovsky hájil. [...] (<i>Nebeský klíč</i> , Okresní muzeum Náchod, Rkps 70, 1763, p. 7) ⁶

The textual comparison of the so-called Keys of Heaven divides the texts into two groups. The first of them contains the original prayers, i.e., those which Martin of Cochem included into his *Golden Key of Heaven*. The other group consists of similar prayers from manuscripts labelled as Key of Heaven (or another variation of this title).

In this paper, we apply two statistical methods to analyze a set of prayers from a particular manuscript. Results obtained strongly indicate that, in spite of some textual variation, the manuscript is an „honest copy” of the original *Golden Key of Heaven* rather than a compilation of various sources, the book by Martin of Cochem being one of them.

2. Language material

The original *Golden Key of Heaven* consists of sixteen chapters of prayers: *Morning and Evening Prayers*, *Prayers in a Church and Prayers at Mass*, *Vespers for the Dead*, *Prayers before and after Confession*, *Prayers before Communion*, *Prayers to the Blessed Sacrament*, *Five Psalms to the Holy Trinity*, *Prayers to Christ the Lord*, *Prayers to the Passion of Christ the Lord*, *Prayers to the Virgin Mary*,

⁵ A Czech translation of Martin of Cochem's *Golden Key of Heaven* from the Orlice Museum in Choceň, Czech Republic.

⁶ A manuscript of *Nebeský klíč* (*Key of Heaven*) from the Regional Museum in Náchod, Czech Republic.

Prayers to the Saints, Prayers for the Annual Feasts, Prayers in General and Particular Needs, Prayers for Various States, Prayers for a Happy Death, Prayers for the Dead. In addition, this huge volume contains numerous discourses on the purgatory, passages instructing readers how to behave when they pray, and, at the end, one finds a story of folk reading with the title *The History of One Monk from Engelland Who Spent Three Days in the Purgatory and Fearful Things He Saw There*. The *Golden Key of Heaven* had been reprinted in this form, regardless of publishers and their locations, until 1770s, when its content was changed by the state censorship.

We compare the abovementioned book with the manuscript *Key of Heaven* found in the Regional Museum in the town of Náchod (hereafter, the text sources are sometimes referred to as *Cochem* and *Náchod*, respectively). The manuscript is markedly thinner than the book.⁷ It begins with *Morning and Evening Prayers* and then continues with *Prayers at Mass, Prayers before Confession, Prayers before Communion, Litany of the Lord Jesus, Devotions at All the Feasts of the Virgin Mary, Litany of Loreto, Tuesday Devotion to Mother Saint Anne, Nine Tuesdays Devotion to Saint Anthony of Padua, Devotion at the Tomb of Saint John, Litany of the Saints, Prayer of Children for Their Parents* and *Prayer of the Passion of Christ the Lord*.

Our previous textological research (Netolická 2015, 2018) shows that the manuscripts contain a common thematic core, such as *Morning and Evening Prayers, Prayers at Mass* and *Devotion to Saint Anthony of Padua*. Other prayers attached to this core display a remarkable variability. The choice of the prayers could be guided by the target audience (prayers for men differed from those for women). Some prayers were shortened; sometimes, a folk story was inserted to the manuscript. Detailed comparisons can provide valuable insights into the process of creating manuscript prayer books.⁸

⁷ While the printed *Golden Key of Heaven* contains more than three hundred prayers, the manuscript *Key of Heaven* consists of about fifty five prayers.

⁸ In our manuscript, such is the case of *Prayer of children for their parents*—except for its end, it was literally copied from *Cochem*. But the *Cochem*'s end of this prayer

A set of ten pairs of prayer texts was used in our analysis. These prayers appear under the same names both in the printed book and in the manuscript. In one case, there are two texts with similar title and similar context in the book, namely *Tuesday Devotion to Mother Saint Anne* and *Tuesday Prayer to Mother Saint Anne* (they are labelled as 07_Kochem_a and 07_Kochem_b). Both of them are included in the sample.

3. Methodology

From the methodological point of view, our corpus is an ideal sample for a quantitative analysis. Literature published at the turn of the 17th and 18th century is generally characterized by its uniformity which was caused by several factors, such as a limited number of authors, the lack of education outside of the Church, censorship, etc. The fact that all texts belong to the same genre, namely prayers, is very important because the homogeneous style enables us to focus on the text features in detail. Thus, we do not have to struggle with a genre-related bias which is often the stumbling block of many quantitative analyses, especially those conducted on large corpora. We work with a small homogenous sample consisting of 20 texts, all of them written in one consistent style.

Given that our aim is to answer the question of whether the *Náchod* manuscript can be considered a copy of the *Cochem* book, we decided to use two quantitative methods from the contemporary stylometry, especially frequency structure indicators. The first method deals with most frequent words (MFW) which is one of the traditional and oldest features in this field. We decided to obtain 100 MFW from the whole corpus and use them to find a similarity among individual texts. From the statistical viewpoint, we chose the cluster analysis as the most common method of data classification. All computations were performed by the R package „Stylo” (Eder et al. 2013). It is necessary to

this prayer was omitted and replaced with this suggestive phrase: *therefore I choose to die rather than make you and them angry*.

mention that we consider word-forms as basic units in the analysis, therefore texts were not lemmatized. Czech is highly inflected, nouns and adjectives are declined into one of seven grammatical cases. The inflectional suffixes indicate also number, gender and animacy. Verbs are conjugated for tense, number and gender. Given that there are several options how to express one meaning, we suppose that the choice of the word-form is a matter of the author's style (the method was successfully applied e.g. to analyze genres in Karel Čapek's writing by Kubát (2016).

Stylometric analyses from recent years prove that n-grams are very powerful features in authorship attribution. An n-gram is a continuous sequence consisting of n units in a given text. N-grams are units which can be used in several fields such as probability, communication theory, natural language processing, computational biology or data compression. In quantitative linguistics, n-grams are usually used in stylometry. The n-grams seem to be appropriate units for this research, considering the high accuracy of authorship attribution (Mikros, Perifanos 2013). Since n-grams can be used with many linguistic units such as word forms, lemmas, syllables, etc., there are several options how to apply n-grams in a stylometry analysis. In this study, we decided to use „the Author's Multilevel N-gram Profile" (AMNP) which is a special method designed for authorship attribution by Mikros, Perifanos (2013). The AMNP combines character and word bigrams and trigrams, namely the 200 most frequent n-grams. Thus, we got 800 features (4×200) for the analysis. The combination of four kinds of n-grams and their connection to language levels is displayed in Figure 1.

Although the analyses based of n-grams are very successful, these artificial units are not convenient from the linguistic point of view, as they may be difficult to interpret (especially character 2- and 3-grams). Nevertheless, AMNP seems to be a suitable tool for our purpose. A supervised classification, namely Support Vector Machine (SVM), was chosen for final statistical assessment. The SVM model enables to classify texts with high accuracy of prediction (more than 90 % in authorship attribution (Mikros, Perifanos 2013).

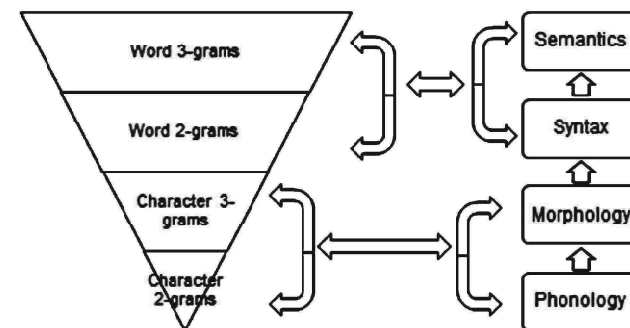


Figure 1. Author's Multilevel N-gram Profile by Mikros (2013)

4. Results

4.1. Most frequent words (MFW) Cluster Analysis

The results of the cluster analysis based on the 100 most frequent words are displayed in Figures 2 a 3. We can see the network in the

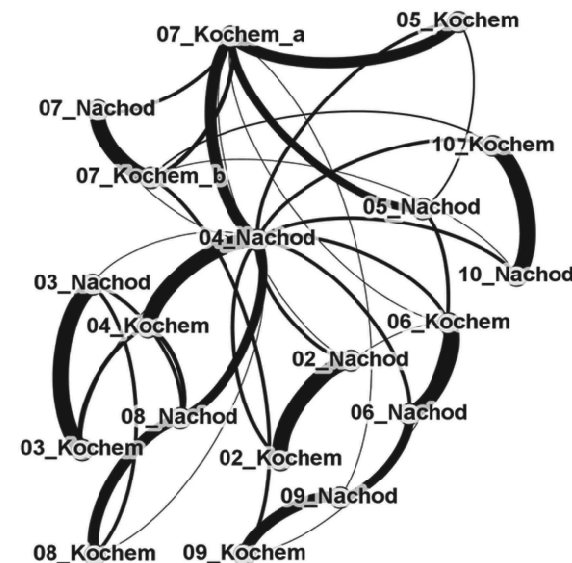


Figure 2. The network of cluster analysis based on 100 MFW

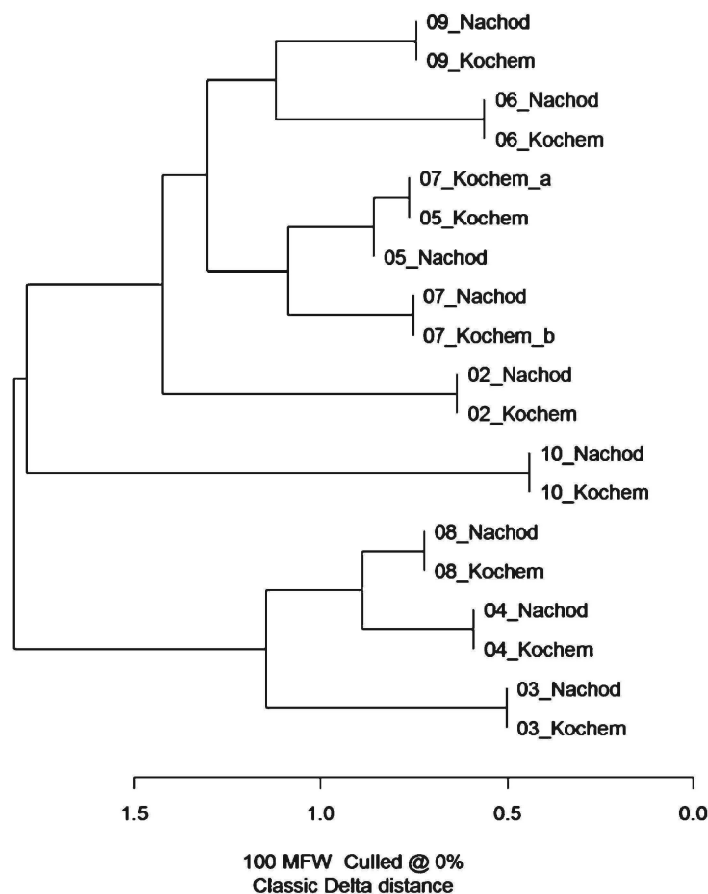


Figure 3. Dendrogram of the cluster analysis using 100 most frequent words.

Figure 2 and the dendrogram in Figure 3. The nodes in the network are connected by edges. The strength of the edges is related to the weights. In other words, the nearest neighbours have the strongest link.

At the first sight we can see strong connections between the pairs of the texts with the same numbers. The majority of them are linked how we expected, and only a few strong connections do not meet our

expectation. To sum up, the cluster analysis indicates that, because of the number of significant links, the *Náchod* is a copy of the *Cochem*.

4.2. Author's Multilevel N-gram Profile (AMNP)

Table 2. The attribution of the texts based on SVM.

Prediction / Reference	<i>Cochem</i>	<i>Náchod</i>
<i>Cochem</i>	10	9
<i>Náchod</i>	0	0

The results displayed in Table 1 show that the SVM model matches up all texts with the *Cochem*. This suggests that all texts of the *Náchod* are copied from the *Cochem*. Taking into account the fact that authorship analyses based on AMNP and SVM can distinguish authors of given texts with more than 90 % accuracy, our obtained results lead to the conclusion that all texts were written by one author. Thus, we can say that, despite some differences, the *Náchod* is very probably a copy of the *Cochem*.

5. Conclusion and discussion

Both methods provide quite a convincing hint that the *Náchod* manuscript is an „honest copy” (with minor differences which can be explained by space restrictions, scribe mistakes, scribe using a dialect of Czech language, etc.) of some chapters from the *Cochem* book. The MFW methods shows very strong links between respective texts, which means that the same words appear in the same parts of the book and of the manuscript. The SVM method classifies all texts as being written by Martin of Cochem.

The methods seem to be applicable to a broader spectrum of questions, such as the opposition between a copy and a compilation of several sources here, the differences between presidents with different political and ideological backgrounds (Kubát et al. 2020), or author-

ship attribution (Mikros, Perifanos 2011). In our context, their applications could lead to a deeper understanding of the phenomenon of re-writing books (like e.g. the development of the ratio of copies and compilations in time, tracing the „geographical trajectories” of similar texts, etc.).

References

- Eder M., 2014, *Stylometry, network analysis and Latin literature*. In: *Digital Humanities 2014: Book of Abstracts*, Lausanne: EPFL-UNIL, pp. 457–458.
- Eder M., Kestemont M., Rybicki J., 2013, *Stylometry with R: a suite of tools*. In: *Digital Humanities 2013: Conference Abstracts*. Lincoln: University of Nebraska-Lincoln, pp. 487–489.
- Eder M., Rybicki J., 2009, *PCA, Delta, JGAAP and Polish poetry of the 16th and the 17th centuries: who wrote the dirty stuff?* In: *Digital Humanities 2009: Conference Abstracts*. College Park Maryland: University of Maryland, pp. 242–244.
- Fidlerová A., Bekešová M., 2007a, *Repertorium rukopisů 17. a 18. století z muzejních sbírek v Čechách II/1 K–L*. Praha: Karolinum.
- Fidlerová A., Bekešová M., 2007b, *Repertorium rukopisů 17. a 18. století z muzejních sbírek v Čechách II/2 M–O*. Praha: Karolinum.
- Flaišman J., Kosák M., 2006, *Editor a text: úvod do praktické textologie*, Praha/Litomyšl: Paseka.
- Kubát M., 2016, *Kvantitativní analýza žánrů*. Ostrava: Ostravská univerzita.
- Kvapil J., 2001, *Ze zahrádky do zahrady aneb od Hortulu animae k Štěpné zahradě Martina z Kochemu. Utváření modlitební knihy barokního typu*, Ústí nad Labem: Univerzita J. E. Purkyně, pp. 42–43.
- Lehár J., 1998, *Česká literatura od počátků k dnešku*, Praha: Lidové noviny.
- Linda J., 2003a, *Repertorium rukopisů 17. a 18. století z muzejních sbírek v Čechách I/1 A–F*, Praha: Karolinum.
- Linda J., 2003b, *Repertorium rukopisů 17. a 18. století z muzejních sbírek v Čechách I/2 H–J*, Praha: Karolinum.
- Martin (von Cochem), Sládek M., Kořínková L., Breň T., 2007, *Veliký život Pána a Spasitele našeho Krista Ježíše a jeho nejsvětější a nejmilejší matky Marie Panny*, Praha: Argo.
- Mikros G. K., Perifanos K., 2011, Authorship identification in large email collections: Experiments using features that belong to different linguistic levels Proceedings of PAN 2011 Lab, Uncovering Plagiarism, Authorship, and Social Software Misuse held in conjunction with the CLEF 2011 Con-

ference on Multilingual and Multimodal Information Access Evaluation, 19–22 September 2011, Amsterdam.

- Mikros G. K., Perifanos K., 2013, *Authorship attribution in Greek tweets using multilevel author's n-gram profiles*. In E. Hovy, V. Markman, C. H. Martell & D. Uthus (eds.), *Papers from the 2013 AAAI Spring Symposium Analyzing Microtext, 25–27 March 2013, Stanford California Palo Alto, California: AAAI Press*, pp. 17–23.
- Mikros G. K., 2009, *Content words in authorship attribution: An evaluation of stylistometric features in a literary corpus*, In: R. Köhler (ed.), *Issues in Quantitative Linguistics*, Lüdenscheid: RAM-Verlag, pp. 61–75.
- Mikros G. K., 2013, *Machine Learning methods in Authorship Attribution and Profiling*. Presentation at IQLA-GIAT Summer School in Quantitative Analysis of Textual Data of Padua 2013.
- Netolická Š., 2015, *Dvojí podoba religiozity na příkladu Zlatého nebeklíče Martina z Kochemu*. In Ferencová et al. *Proměny konfesijní kultury v Evropě*. Olomouc: Filozofická fakulta Univerzity Palackého, pp. 227–235.
- Netolická Š., 2018, *Zlatý nebeklíč Martina z Kochemu a rukopisné modlitební nebeklíče ve světle osvětské cenzury*. In Tichý, M. – Sichálek (eds.) *Mezi kritikou a poezií. Ladislavu Soldánovi k osmdesátinám*. Opava, Slezská univerzita v Opavě.
- R Core Team, 2013, *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Online: <http://www.R-project.org>.
- Rybicki J., 2012, *The great mystery of the (almost) invisible translator: stylometry in translation*. In: M. Oakley and M. Ji (eds.), *Quantitative Methods in Corpus-Based Translation Studies*. Amsterdam: John Benjamins, pp. 231–248.
- Rybicki J., Heydel M., 2013, *The stylistics and stylometry of collaborative translation: Woolf's 'Night and Day' in Polish*. „Literary and Linguistic Computing” 28(4), pp. 708–717.
- Vašica J., 1995, *České literární baroko*, Brno: Atlantis.