

## **Naming and Renaming Texts.**

### **Distant Reading of Middle High German Rubrics in Miscellany Manuscripts<sup>1</sup>**

Dr. Gustavo Fernández Riva

In my presentation today I will talk about rubrics in Middle High German miscellany manuscripts. I will try to show how techniques from what is usually called “distant reading” can help us analyze this material and gain insight into MHG literature and how medieval scribes and readers interpreted literary works. My exposition will have the following parts:

- 1. The canon of Middle High German Literature as Network .** I will introduce the theoretical foundation of the research.
- 2. The MHG short texts in rhyming couplets .** Where I will talk about the texts that compose MHG miscellany manuscripts and their features. Then I will explain the more concrete corpus used, the
- 3. Rubrics in Miscellany Manuscripts .** Then the central part, the
- 4. Analysis of Rubrics in the Corpus** followed by some reflexions on
- 5. Rubrics and Genre .**
- 6. Final Thoughts .**

#### **1. The canon of Middle High German Literature as Network**

I would like to start from an abstract perspective, thinking about how different conceptions of literature, and of literary canons, can be represented as different structures. For example, in his famous books *The Western Canon* and *Anxiety of Influences* Harold Bloom considers the canon as a chain of authors who influence each other and could be graphically represented as a directed graph (image 1). However, usually a canon is thought of as a list containing a limited number of items, works that should or at least deserve to be read (image 2). This listing can be also found in Bloom's *Western Canon*, as shown in the image. Nonetheless, this two representations don't show one of the most relevant feature of a literary canon: that it leaves things outside its scope. A canon is by definition a subset of all possible texts that correspond to some parameters. In this sense, it could be represented as a Venn diagram, where the canonical texts are only a small subset of the whole set (image 3). And this image is misleading, because the canon set would be a lot smaller, almost invisible.

---

<sup>1</sup> This is the draft that guided the presentation at the conference *The Medieval Canon in the Digital Age*, Ghent (Belgium), 17<sup>th</sup> September 2018. This is a draft, not to be considered a definitive version. The author holds the copyright of this and all future versions.

### Exemplary Representation of Bloom's Conception of Canon

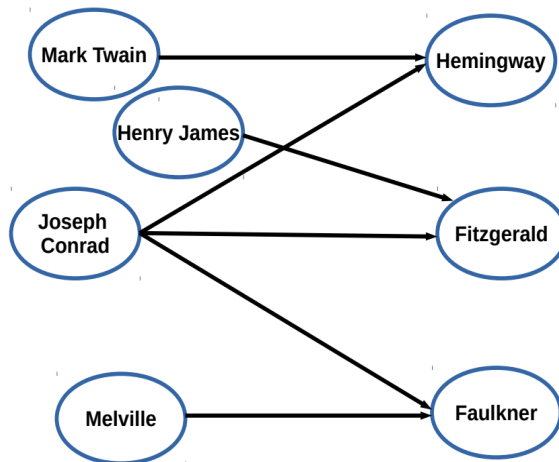


Image 1

<b>Isaac Bashevis Singer</b> <i>Collected Stories</i> <i>In My Father's Court</i> <i>The Manor, The Estate, The Family Moskat</i> <i>Satan in Goray</i>	<b>ARABIC</b> <b>Najib Mahfuz</b> <i>Midaq Alley</i> <i>Fountain and Tomb</i> <i>Miramar</i>
<b>HEBREW</b> <b>Hayyim Nahman Bialik</b> <i>Shirot Bialik: The Epic Poems</i> <b>S. Y. Agnon</b> <i>In the Heart of the Seas</i> <i>Twenty-one Stories</i>	<b>Adunis</b> <i>Selected Poems</i> <b>Mahmud Darwish</b> <i>The Music of Human Flesh</i> <b>Taha Husayn</b> <i>An Egyptian Childhood</i>
<b>Aharon Appelfeld</b> <i>The Immortal Bartfuss</i> <i>Badenheim 1939</i> <b>Yaakov Shabtai</b> <i>Past Continuous</i> <b>Yehuda Amichai</b> <i>Selected Poetry</i> , translated by Stephen Mitchell and Chana Bloch <i>Travels</i> , translated by Ruth Nevo	<b>LATIN AMERICA</b> <b>Rubén Dário</b> <i>Selected Poetry</i> <b>Jorge Luis Borges</b> <i>The Aleph and Other Stories</i> <i>Dreamtigers (The Maker)</i> <i>Ficciones</i> <i>Labyrinths</i> <i>A Personal Anthology</i> <b>Alejo Carpentier</b> <i>Explosion in a Cathedral</i>

Image 2

### The Literary Canon as a Venn Diagram

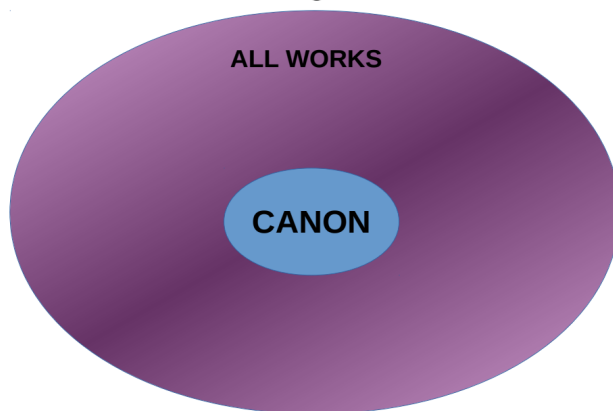
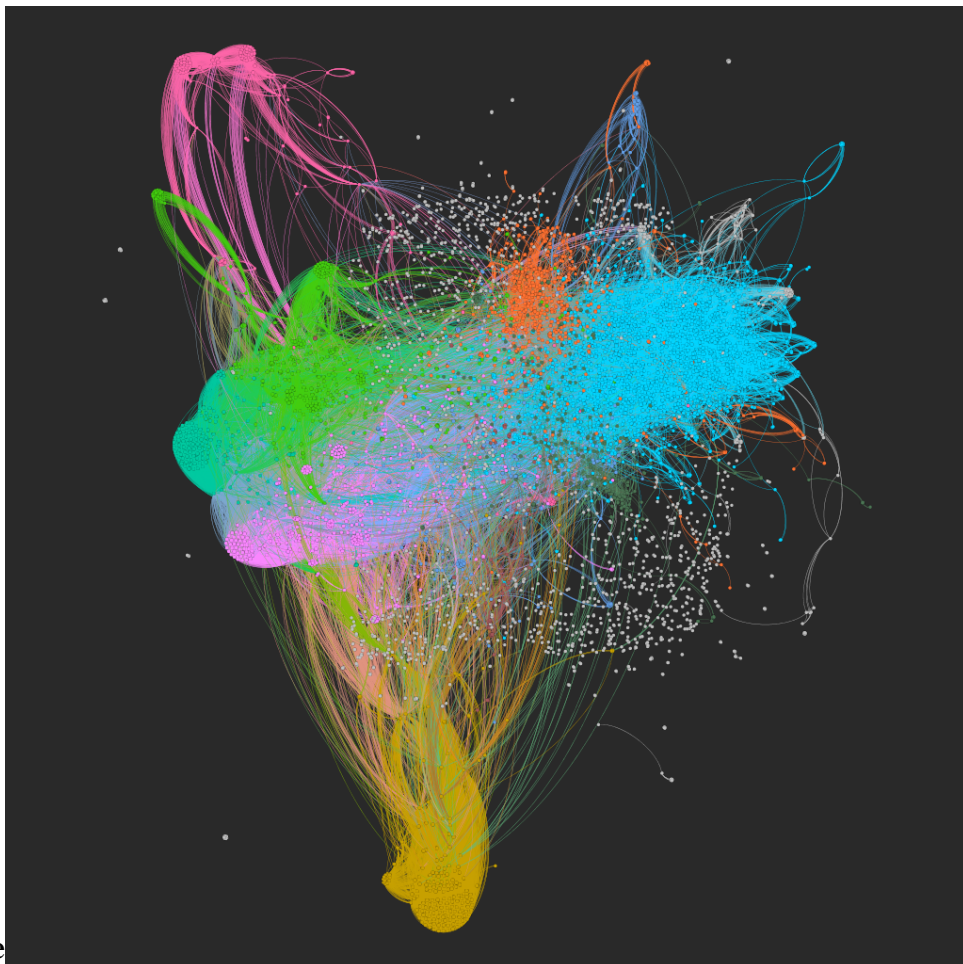


Image 3

The need to select a group of canonical texts is understandable, specially when it comes to aesthetic objects. A reader might want to focus on what are considered the best works available. The need to create canonical lists is specially true of school and university curricula; but we can also understand it as a more general tendency. For example, in the contemporary world the internet is overpopulated by ranked lists of things, and literature is no exception. This shows that literary canons do not only exist for practical reasons (i.e. limited time and need to select only the best things); they have a more fundamental reason to exist: we humans enjoy organizing things into lists, and something that is the first in a ranking seems more appealing.

However, as literary scholars we should consider the problems and biases that come with our tendency to create literary canons. To truly understand how literature worked in a given context, we need to see beyond the canon.

One of the ways the Digital Humanities promise to open up this non-canonical perspectives consists in offering new ways of dealing with big quantity of texts; quantities that are not so easily analyzed with traditional philological methods. In the digital age we should base our understanding of literature on a different paradigm than the ones that support canon constructions. Instead of thinking of literature as a a directed graph, a list or a Venn diagram, we should consider it a network. Of course, the idea of the literary network long precedes the digital age, but with the aid of Digital Humanities this idea gain a renewed relevance. A network of literature can be constructed in many different ways according to different parameters. In figure 4, I have created a possible network of Middle High German literature based on its manuscript tradition using the database of the online catalogue *Handschriftencensus*. Each node represents a text and the edges connect works when they share a manuscript. The colors of the graph were automatically generated by *Gephi* with the modularity function, which uses a community detection algorithm to highlight the interrelated nodes, which, as it turns out, mostly correspond to different genres. There are multiple other possible networks for a given literary system. This one highlights how mixed and interrelated the Middle high German literature is in the documents that have come down to us. Today I will focus on a subset of this network of Middle High German literature: the short poetry composed in rhyming couplets.



2. The

Image 4

## MHG short texts in rhyming couplets

Usually only between 300 and 600 verses. They can be narrative or not and have many different subjects and purposes. They are to be found in miscellany manuscripts that compile lots of them, sometimes more than 200. The corpus I will be working with contains over 800 different works, which I have divided and tagged according to traditional modern genre categories:

- *Minnereden* (not narrative secular) - 273
- *Maeren* (narrative secular) - 231
- *Fables* (narrative animal characters) - 83
- Religious / *Moral* / *Didaktik* (saint's lives, marian miracles and other not narrative religious, moral or didactic poetry) -167
- Other - 75

Of course this division could be more detailed. It is particularly unfair to religious literature, as it does not make a more granular distinction, but for my purposes in this presentation, it is good enough.

A feature of these sources is that they have not been traditionally part of the Middle High German literary canon. The most canonical works from medieval Germany are the epic poems of around 1200: *Nibelungenlied*, Wolfram von Eschenbach's *Parzival*, Gottfried von Strassburg's *Tristan*; and also the lyrical poetry of Walther von der Vogelweide. However, they have been in some periods subject to an analytical method I would call "Systemic Reading". What I mean under Systemic Reading are works that are located at the crossroads of a catalogue and a history of literature. Their main objective is not interpretation of individual works, but to create an ordered typology and understand the relationships within a complex literary system composed by a considerable number of texts. In only two years, between 1967 and 1968 three important books with this perspective appeared:

Mihm, Arend. *Überlieferung und Verbreitung der Märendichtung im Spätmittelalter*. Heidelberg: Winter, 1967.

Fischer, Hanns. *Studien zur deutschen Märendichtung*. Tübingen: Max Niemeyer, 1968.

Brandis, Tilo. *Mittelhochdeutsche, mittelniederdeutsche und mittelniederländische Minnereden; Verzeichnis der Handschriften und Drucke*. Munich: Beck, 1968.

Afterwards, many other books addressed the issues presented by this corpus. One of their key concerns has been differentiating sub-genres and their features.<sup>2</sup> However, in the last decade we've

2 Hans-Joachim Ziegeler, *Erzählen im Spätmittelalter: Mären im Kontext von Minnereden, Bispeln und Romanen*, Münchener Texte und Untersuchungen zur deutschen Literatur des Mittelalters 87 (München: Artemis, 1985); Werner Röcke, *Die Freude am Bösen. Studien zu einer Poetik des deutschen Schwankromans im Spätmittelalter* (München: Wilhelm Fink, 1987); Joachim Heinze, "Märebegriff und Novellentheorie. Überlegungen zur Gattungbestimmung der mittelhochdeutschen Kleinepik", *Zeitschrift für deutsches Altertum und deutsche Literatur* 107 (1978): 121–38; Walter Haug, "Entwurf zu einer Theorie der mittelalterlichen Kurzerzählung", in *Brechungen auf dem Weg zur Individualität. Kleine Schriften zur Literatur des Mittelalters* (Tübingen: Max Niemeyer, 1995); Ingrid Strasser, *Vornovellistisches Erzählen: mittelhochdeutsche Mären bis zur Mitte des 14. Jahrhunderts und altfranzösische Fabliaux* (Wien: Fassbaender, 1989); Franz-Josef Helznagel, "Verserzählung - Rede - Bispeln. Zur

seen a renewed interest in cataloguing and ordering the texts as a more efficient way of understanding them. There are two books in this perspective:

Klingner, Jacob, and Ludger Lieb. *Handbuch Minnereden*. Berlin, Boston: De Gruyter, 2013.

Eichenberger, Nicole. *Geistliches Erzählen: Zur deutschsprachigen religiösen Kleineliteratur des Mittelalters*. Berlin: De Gruyter, 2015.

However, this systemic approach can't be considered Distant Reading strictly speaking. Although it uses some statistical methods, it most commonly collects and draws conclusions from a lot of smaller close reading analysis. Against the background of this scholarly tradition, I don't think Digital Humanities and Distant Reading is radically new. It is a way of pursuing a trend that is already present in the field, but with different strategies and tools, that hopefully can improve the results of previous approaches.

### 3. Rubrics in Miscellany Manuscripts

Why rubrics? We still don't have enough Middle High German texts digitised in a proper manner to perform distant reading analyses. However, rubrics are relatively easy to collect and digitize, and could grant us some sort of insight into MHG texts, apart from the fact that they are important sources in themselves if we want to understand how medieval readers thought about literature. In this sense, rubrics are similar to modern titles, which according to Franco Moretti, help us “catch a glimpse of the literary field as a whole” when the texts themselves are not available.<sup>3</sup> However, medieval rubrics can only partially be compared to modern titles.

One of the few papers on MHG rubrics in miscellany manuscripts was written by two scholars from the University of Vienna in 2017.<sup>4</sup> In this article they identify three functions for rubrics in the sources:

- 1- Separate;
- 2- Indicate a text's content;
- 3- Indexing of a story.

In their article they also explain that:

"Our initial hypothesis was that the use of titles for short verse narratives developed gradually from the thirteenth to the fifteenth century, starting with non-titles and ending

---

Typologie kleinerer Reimpaardichtungen des 13. Jahrhunderts”, en *Eine Epoche im Umbruch. Volkssprachliche Literalität 1200-1300. Cambridge Symposium 2001*, ed. Christa Bertelsmeier-Kierst y Christopher Young (Tübingen: Max Niemeyer, 2003), 291–306; Klaus Grubmüller, *Die Ordnung, der Witz und das Chaos. Eine geschichte der europäischen Novellistik im Mittelalter: Fabliau - Märe - Novelle* (Tubinga: Niemeyer, 2006); Klaus Grubmüller, “Mittelalterliche Novellistik im europäischen Kontext: Die komparatistische Perspektive”, en *Mittelalterliche Novellistik im europäischen Kontext*, ed. Mark Chinca, Timo Reuvekamp-Felber, y Christopher Young, Beihefte zur Zeitschrift für deutsche Philologie 13 (Berlin: Erich Schmidt, 2006), 1–23.

3 Franco Moretti, “Style, Inc. Reflections on Seven Thousand Titles (British Novels, 1740-1850)”, *Critical Inquiry* 36, núm. 1 (2009): 134.

4 Matthias Meyer y Nicola Zotz, “How to Name a Story? Rubrics – Headings – Titles”, en *The Dynamics of the Medieval Manuscript: Text Collections from a European Perspective*, ed. Karen Pratt et al. (Göttingen: Vandenhoeck & Ruprecht, 2017), 203–16.

with specific titles. However, we were wrong. Medieval manuscripts are never predictable and no development is linear."

In this presentation one of my goals is to explore and assess if MHG rubrics in miscellany manuscripts are actually unpredictable and non-linear as these scholars suggest.

#### 4. Analysis of Rubrics in the Corpus.

The corpus of rubrics was made using mostly old library catalogues and manuscript facsimiles. I transcribed almost 1800 rubrics from 40 manuscripts, which correspond to over 800 different works. It still needs some work and when it is more complete I could redo the analysis, but I think what I have now is representative.<sup>5</sup>

The first very general inquiry we can make to this corpus is asking how are these rubrics constructed. One first approach would be to compare them to the modern titles for the same texts. The most obvious comparison feature is **length**. The average number of words in **modern titles** in the corpus is **3.17**, while the average number of words in **medieval rubrics** is **7.88**. This suggests a very strong difference. Rubrics offer more information. If one disaggregates the data for each manuscript and plots the average length of titles according to the year the manuscript was produced, one sees that there is a very slight tendency to increase the length (image 5). However, this trend is very small in comparison with the more relevant fact, that there is a very strong variation in almost all periods. Another relevant fact is that after 1500 all manuscripts have long titles, which one could argue is because of the influence of the printing press and its famously long titles.<sup>6</sup>

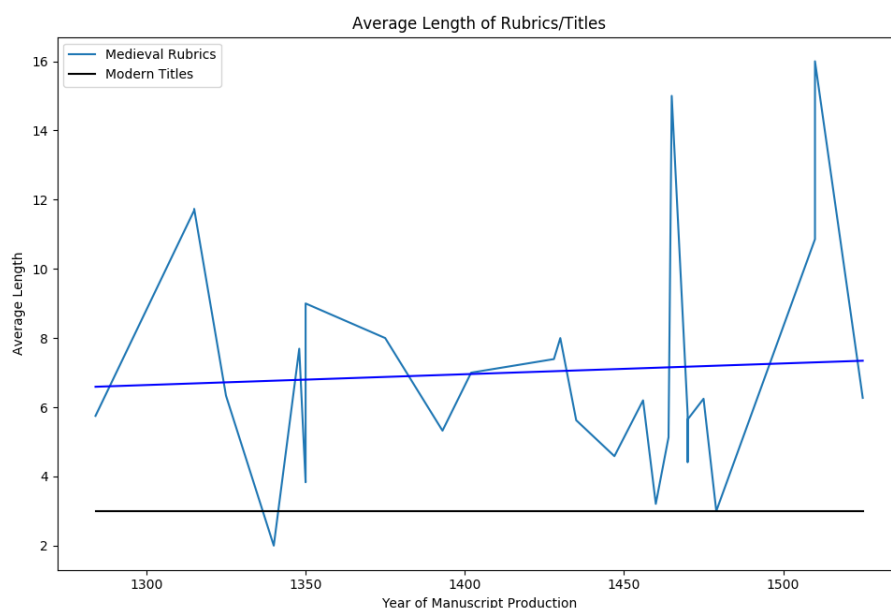


Image 5

<sup>5</sup> The raw data can be found at: [https://github.com/GusRiva/projects\\_data/tree/master/rubrics](https://github.com/GusRiva/projects_data/tree/master/rubrics)

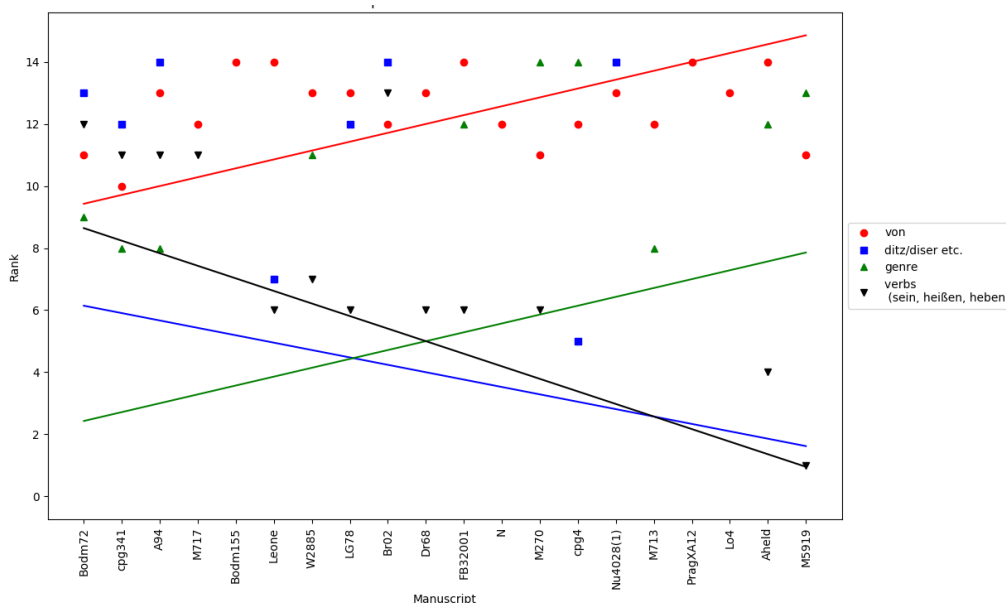
<sup>6</sup> Best fit line composed using the "Least squares method".

Now we could ask ourselves how medieval rubrics in these miscellany manuscripts are constructed in order to better understand their function. If we look at the most frequent words in the entire corpus, for example in a word cloud we get a pretty accurate idea of what they are like (image 6).<sup>7</sup>



*Image 6*

With only the most frequent words we can compose the beginning of a typical rubric: *ditz ist daz mere...* (this is the story...). However, some early manuscripts with more works have too much weight in this word cloud. We can disaggregate the information according to manuscripts. For that, I ranked the top words for each manuscripts. There are many function words, like articles that are very common, which corresponds to the normal proportion in any given sample of the language. Many other words appear just one or two times. However, there are some prominent words that occur above average that can give us valuable information about this genre. To better analyze this, I constructed the following plot.



*Image 7*

7 Word cloud based on the 55 most frequent words in the corpus. The 5 most frequent are *sîn* (408); *diser* (376); *daz* (186); *mere* (157); *hie* (112).

There are four categories I found interesting to plot: 1- the word *von*, 2- demonstrative pronouns *ditz* and *diser*, 3- words that refer to the text (“genre”), 4-copulative verbs and other verbs that refer to the function of rubrics. In the x axis we have the manuscripts ordered from oldest to newest, in the y axis we have the rank of the words within the rubrics of the manuscript. 5 is the most frequent word, 1 the fifth most frequent word. And I am considering all words, there is no stop words list. These common words and their evolution can tell us a lot about the corpus. Let's analyze them one by one, according to how high their rank is.

1- The word *von*. This preposition is used to indicate the subject of the text and could be translated as *of* or *about*, like the *de* so common in latin headings. Some rubrics have an introductory statement like *ditz ist von*, but many other just offer a prepositional phrase like ***von dem ritte mit der halben bir*** (About the knight with the half pear) or ***von dem armen Heinrich*** (About the poor Henry). This is very similar to a modern title, which are most often than not, composed of a nominal phrase. However, for some reason scribes considered it worth while keeping the preposition, making it explicit that this rubric describes what the text is about. This property of the rubrics shows that they are constructed in such a way that they don't just have the function of describing the content of the following text, but that they explicitly state that function.

2- In lots of cases there is a mention of the “genre”. I will come back to this in a moment.

3 – *ditz/disser* is a demonstrative that works almost as a deictic marker: this text in front of you; this thing you are listening to. This word link, on the one hand, the rubric to the text and, on the other, the reader to the materiality of the codex. The convention of modern titles don't usually have this kind of devices pointing to the materiality of the book. But the medieval rubrics are constructed as if they were directly addressed to the recipient, separated from the text and pointing at it. We understand modern titles in the same way, but they don't make that explicit.

4 - Finally, there are some verbs that refer to the function of the rubrics. It is common to encounter *sîn* (to be), usually combined with *von* or something similar (this is about etc.). Not as common, but also normal is *heissen* (is called). Finally, there is also the expression *Hier hebt sich an* which could be loosely translated as “here begins”, literally “here X raises itself”. In this case the expression has a very strong demonstrative value. The *hier* refers to the semantic universe of space and the page more than a performance situation, which could be implied with *jetz* (now). In this case, the words that in many stories are at the beginning of the text and indicate the start of the narrative are dislocated to the rubric, which acquires the function of introducing the tale.

It is not uncommon to find many of these kinds of words together in the same rubric in constructions such as *ditz spruch ist von..* (this discourse is about); *hie hebt sich an daz maere von...*(here begins the story of)



The terms I discussed above are, for the most part, function words. An interesting question would be what are the other most common adjectives and nouns:

*vrouwe* (93); *got* (61); *schoene* (60); *ritter* (59); *guot* (58)

Interestingly women are the most predominant subject. This is because of the *Minnereden*, where love and women are usually topics, but also religious poetry for Maria. Knights also figure prominently. God for the amount of religious literature. But one of the most interesting words is *schoene* (beautiful). This word can be used in two different ways. There are titles like *von ein schoene vrouwe* (about a beautiful woman); *ditz ist von der gans / daz was ein schoene jyncvrowelin* (this is about a goose, that was a **beautiful** maiden). Here *schoen* is used to describe characters or events in the text. However, more common is the expression *ditz ist ein schoene maere* (This a **beautiful** maere). Like *Ditz ist ein schones mere / von einem ritter lobere* (This is a **beautiful** maere / about a praiseworthy knight). In this sense, like in the words we analyzed above, the terms don't really refer to the content of the text. In this case the adjective highlights the function of the rubric as an instance of evaluation of the story. The rubric is the place where the texts is judged and valued. Like the demonstrative pronoun *ditz*, the *schoen* increases the distance between rubric and the text; as if the rubric was part of a different voice presenting the text. And I think this makes sense. In modern literature, title and text are the creation of the same author, the title is part of the text. The communicative act between author and reader starts with the title. In the medieval rubrics on the contrary, they are not part of that communicative act between author and receptor. In the rubric, the scribe as the intermediary between the author and the reader expresses themselves. It is the place where they present the text and evaluate it. For that reason, rubrics are valuable, because they let us see the signature of the scribes, those we only rarely express themselves directly.

## Genre.

Lets go back to the words for genre. The common words we can consider are: *mere*, *spruch*, *rede*, *buoch*. What do they mean?

**Maere** is a very common MHG word. Means narrative, news, story, something that someone tells. Originally it had the meaning of “well known/famous” and derived to something someone tells that is worth telling.

**Spruch** is related to *sprechen*, speak. It means something that is “said”.

**Rede** is very similar. Related to the verb *reden*, that also means to speak or talk. However, it can be used to define a language and the general ability to utter speech. *Rede* comes from words meaning to account for something and to reason.

**Buoch** , same root as english “book”. They tend to refer to written texts. Book is always related to

written texts and not to orality as the previous terms.

All these terms are generally used in the rubrics in my corpus to categorize the texts. The question I would like to ask is whether there is a correlation between those names and the modern generic categories. The first thing I did was to plot exactly that question: how do the medieval terms and the modern genre classifications correlate. In the following image we see how many texts of each of the modern genres is presented in the manuscripts using what medieval terms.

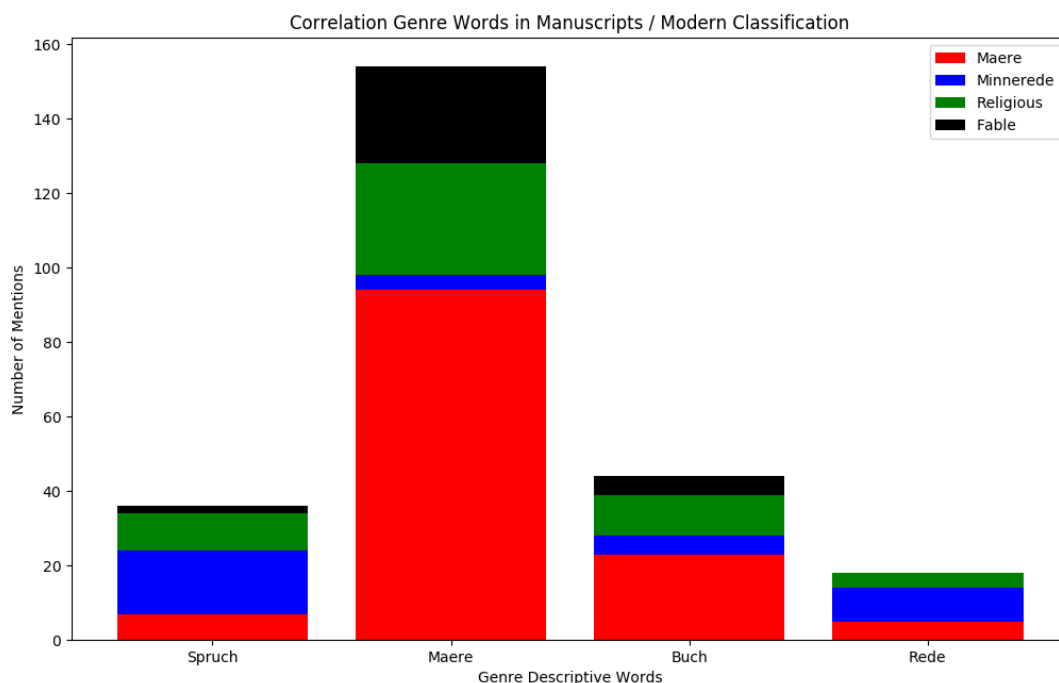


Image 8

At first sight there seems to be some sort of correlation. Mostly *maeren* and fables are mentioned as *maere*. *spruch* and *rede* are used in the manuscript for the modern genre we call *Minnereden*. Religious texts are evenly distributed. However, if we disaggregate the information according to manuscript, we get a more precise idea of why this happens. The plot in image 9 shows that the word *maere* appears mostly in the earliest manuscripts, while the word *Spruch* appears mostly in the oldest ones. *Buch* and *Rede* have way fewer occurrences. If we zoom (image 10) in the graph we can have a better picture of these less common terms: *Buoch* is mostly evenly distributed, while *rede* also appears stronger in the late manuscripts. So, actually we might explain the use of the terms with different moments and fashions and not by the features of the texts.

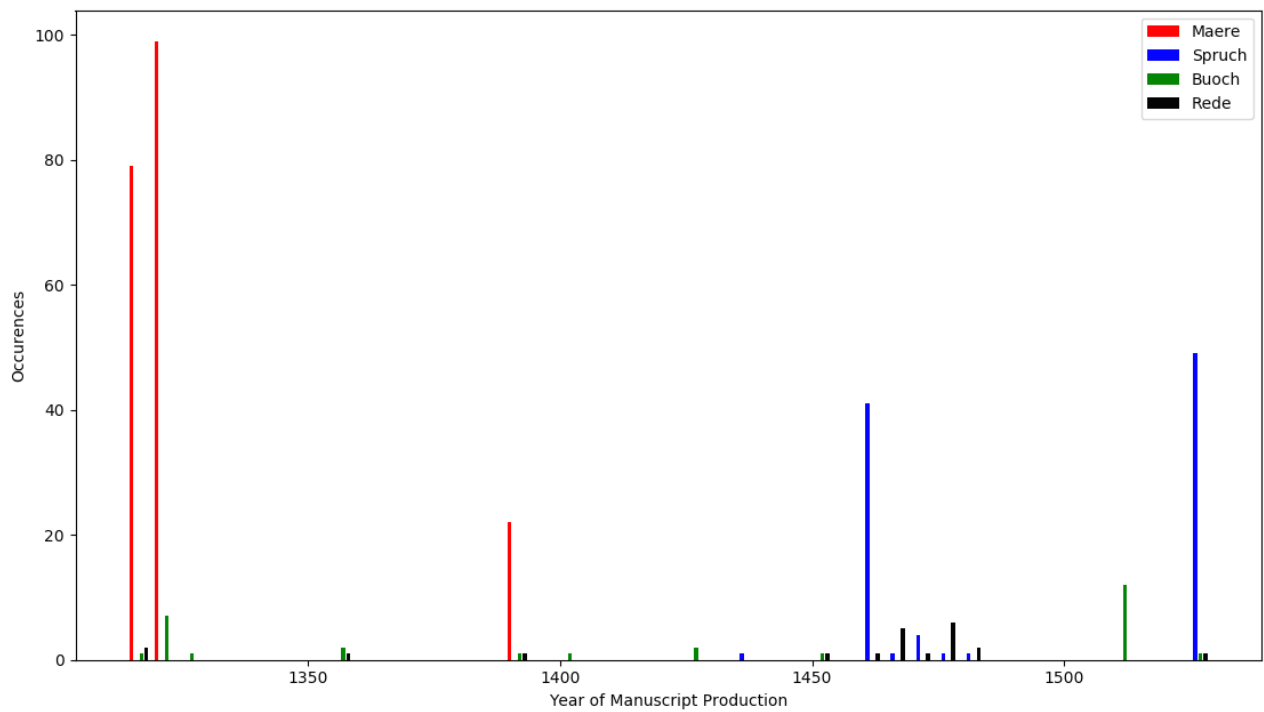


Image 9

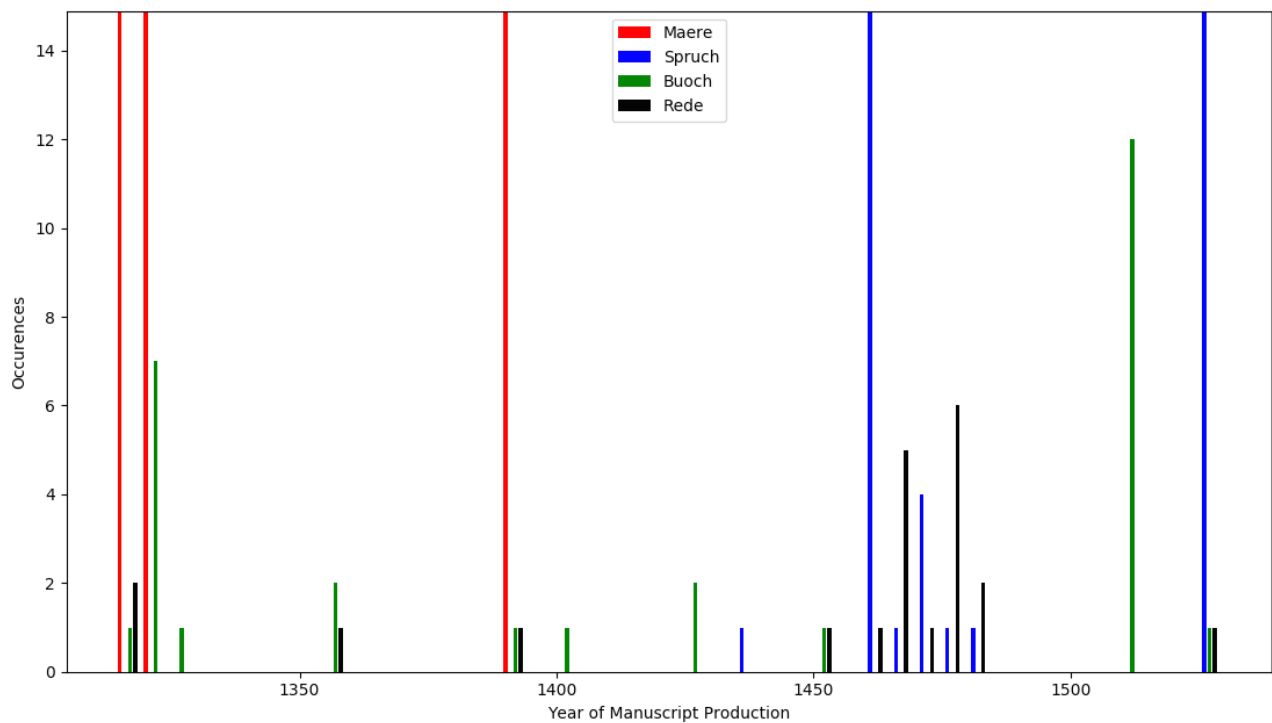


Image 10

If this hypothesis is true, then earlier manuscripts, that use predominantly *maere*, should have more *maeren* according to the modern use, and late manuscripts, which use predominantly the word *spruch*, and *buoch* more *minnereden*. That is in fact the case, as shown in image 11:

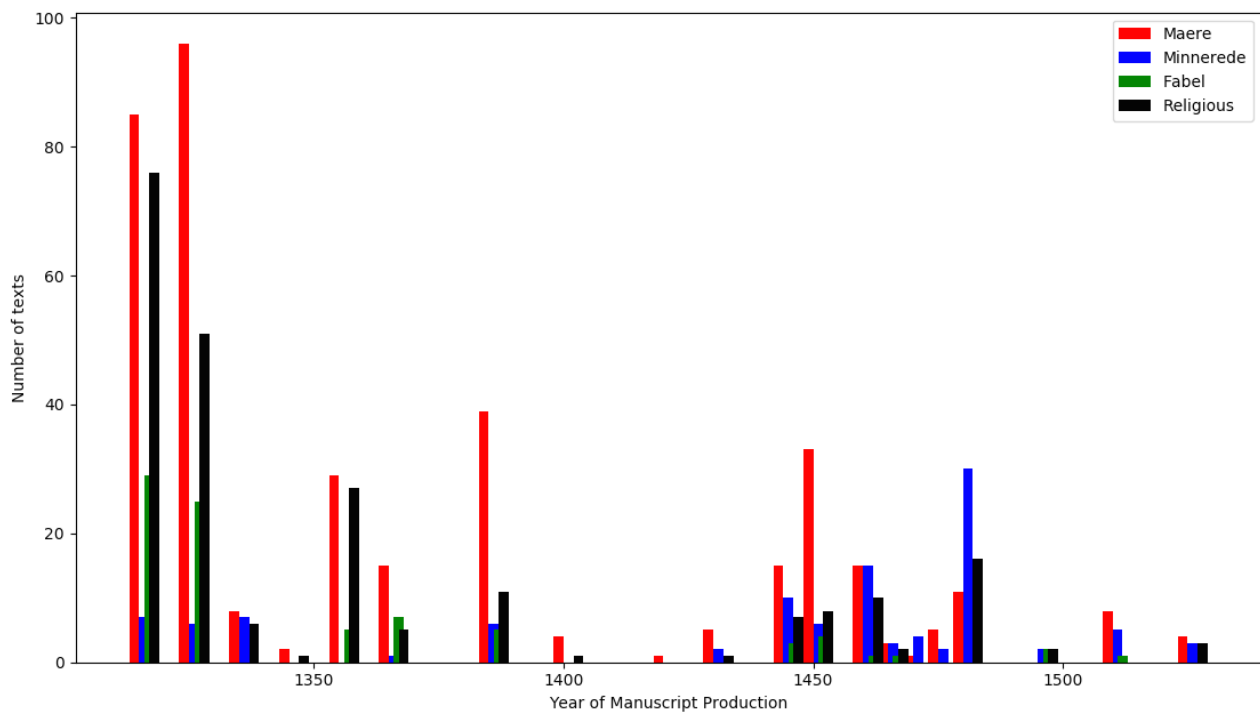


Image 11

We can conclude that the use of *maere*, *spruch*, *buoch* and *rede* depends more on the time of production of the manuscript than on the features of the text. Finally, if we apply most common words to medieval terms and to modern titles, the conclusion that medieval terms don't correlate to content gains even more support:

For medieval terms:

Mere: schoene (27); got (14); ritter (14); guot (13); strickære (11)  
 Buoch: vrouwe (10); got (9); manec (6); lesen (4); helpen (3)  
 Rede: schoene (4); graserinne (2); liebe (2); lieben (2); lîcham (2)  
 Spruch: guot (6); hōvesch (5); leinlin (4); louf (4); künic (3)

For modern genres:

Maere: ritter (38); got (28); künic (27); schoene (25); vil (22)  
 Minnerede: vrouwe (31); guot (16); minnen (16); schoene (11); van (10)  
 Fable: wolfe (19); got (9); strickære (9); lëren (7); katze (4)  
 Religious: vrouwe (17); guot (15); got (10); vil (9); wîp (9)

There is no clear principle explaining the predominance of certain words for the medieval terms, as far as I can see, only the name of the author *Der Stricker* and the adjective *schoen* is used for *Maeren*. However, when divided according to modern genres, we see that words in the rubrics tend to correlate more with the subject. *Maeren* have knights and kings. *Minnereden* have women and love. *Fables* have wolves and cats. Religious texts also have women (Maria) and God. So, this seems to confirm that modern critics created genres according to similarities in topics, while medieval terms don't follow the subject, but are a matter of time and fashion.

## 6. Final thoughts

I have tried to explore timidly how a distant reading of MHG rubrics can offer insight into the network of medieval short literary texts in miscellany manuscripts. However, I need to mention that the digital age is also appropriate for the opposite approach, the one focusing on materiality and the analysis of a particular source. While digital corpora and digital editions are still rare for MHG, the digitization of manuscripts, where Germany is really advanced, permits to explore documents much more easily than before.

For example, if we stay in the field of rubrics, we could analyze the use of rubrication in specific manuscripts and consider, not only their text, but also their material features. One example of the richness in the materiality of a particular heading can be found in fol. 4R of Heidelberg, Universitätsbibl., Cpg 4; the heading of *Rudolf von Ems* by Willehalm von Orleans (image 12). There would be a lot to say about this image from a close reading perspective. But another time.

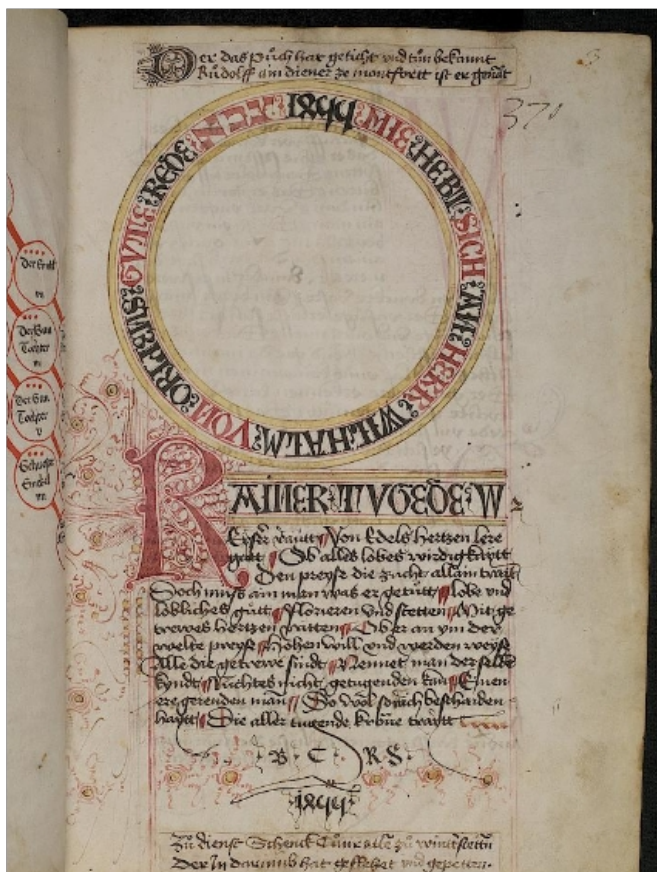


Image 12