

Quantifying Medieval Scribal Habits – The Case of Abbreviations in West Norse Manuscripts

Eva Pettersson^a, Lasse Mårtensson^b and Veturliði Óskarsson^b

Department of Linguistics and Philology, Uppsala University

Department of Scandinavian Languages, Uppsala University

Keywords: West Norse manuscripts, abbreviations, NLP, Old Icelandic, Old Norwegian

Abbreviations were an integral part of the medieval script with Latin letters, as a way of increasing the writing speed and saving parchment. Originating in Latin texts, the system was transferred into texts in vernacular languages, although to a very varying degree. In the Nordic medieval vernaculars, the West Norse area (Norway and Iceland) made a greater use of abbreviations than the East Norse area (Sweden and Denmark). Furthermore, within the West Norse area, the Icelandic manuscripts are especially characterized by an elaborated use of abbreviations. In the *Menota Handbook*, ch. 6.1, it is stated that as much as one third of the words are abbreviated in some Icelandic manuscripts.

In the present study, the West Norse abbreviation system is in focus. Our aim is to extract as much information as possible regarding the use of the abbreviations from digital editions of Old West Norse texts, by combining digital methods for information extraction and qualitative analysis of the retrieved data. We have divided our investigation into four separate, but interrelated, questions:

- Does *the frequency with which abbreviations are used* increase or decrease during the Middle Ages in Old Icelandic manuscripts, and if so to what degree?
- Does *the number of different abbreviations* increase or decrease during the Middle Ages in Old Icelandic manuscripts?
- Are there any common features of the words that are abbreviated and not abbreviated respectively, regarding frequency or linguistic characteristics?
- What differences in the frequency with which abbreviations are used can be observed between the Old Icelandic and the Old Norwegian manuscripts?

For our digital method in question 2 above, we need digital texts transcribed in a format where both abbreviation sign and expansion are available. Texts having only the expansion can also be used for questions 1, 3 and 4. Furthermore, we need texts of a certain length (at least 10 000 tokens) in order to make the quantification reliable. We have used all available texts meeting the demands above in the digital text archives in *Menota*¹ and *Emroon*². Within this text corpus, we have texts from both medieval Norway and Iceland, allowing for a comparison between Iceland and Norway, and texts from different times, allowing for the investigation of the chronological development. In total, the Icelandic dataset contains approximately 350,000 words from the time period 1280–1425, and the Norwegian dataset

¹ <https://clarino.uib.no/menota/catalogue>

² <https://www.emroon.no/#>

consists of roughly 720,000 words from the time period 1200–1350. Of course, when more material becomes available, the picture given by our investigation will be further nuanced. A common formal classification of the medieval abbreviations is that of dividing them into suspensions (the end of the word is abbreviated), contractions (the middle of the word is abbreviated), superscript letters and special signs (e.g. Hreinn Benediktsson 1965: 85).

Regarding the content of the abbreviations, i.e. what linguistic units they represent, they can be divided into two types: 1) abbreviations representing lexical or grammatical forms and 2) abbreviations representing graphemic-phonetic units. Many of the words that are abbreviated in a lexical way are highly frequent words, either names, highly frequent within a certain text (e.g. *Gunnlaugr in Gunnlaugs saga*), or other words recurring in many texts (e.g. verbs like *segja* or *méla*). It should be noted that one and the same abbreviation sign can be used for both lexical and graphemic-phonetic representation; superscript ‘m’ can be used in lexical abbreviations like ‘r’ + superscript ‘m’ for *riddurum* (dat. of *riddari*) as well as for the sequence ‘um’, occurring in different words. It is also known that recurring formulas are often abbreviated through suspensions, e.g. in manuscripts with poetry or legal texts (e.g. Hreinn Benediktsson 1965: 87). In our investigation, we will address both these perspectives; we will account for the type of abbreviation as well as for the units being abbreviated.

Indeed, for all the research questions above, tentative answers can be given beforehand. It is sometimes stated in the literature that the earliest Icelandic manuscripts are less abbreviated than the later ones, and one could assume that there is a general trend during the Middle Ages towards a more frequent use of abbreviations (nr 1). Also, one could assume that the number of different abbreviations would decrease during the Middle Ages, parallel to the reduction of the number of different ordinary letter forms (nr 2). Furthermore, it is very likely that the frequency of the words affects the proneness of the scribe to abbreviate (nr 3), and finally it is often pointed out that the tendency to abbreviate is stronger on Iceland than in Norway (4). However, our empirical basis and our digital methodology will allow us to quantify the use of abbreviations in West Norse manuscripts in a way that has not been done before.

In our study, we show that some of the previously stated hypotheses about the use of abbreviations in West-Norse manuscripts could be validated on an empirical basis. Abbreviations appear to be used gradually more frequently during the Middle Ages on Iceland. The number of different abbreviation signs do not increase in the texts that we have investigated, however; on the contrary, certain abbreviation signs used in the older texts have disappeared in the later text. The tendency that Icelandic manuscripts are more heavily abbreviated than the Norwegian ones, is also confirmed and quantified in our investigation.

References

Hreinn Benediktsson (1965): Early Icelandic script as illustrated in vernacular texts from the twelfth and thirteenth centuries. Íslenzk handrit. Series in folio 2. Reykjavík.

The Menota Handbook v. 3.0. Odd Einar Haugen et al.

https://www.menota.org/HB3_index.xml