Computing Medieval Primary Sources from the Vadstena Monastery: Arguments for the Primary Source Text

Karl G. Johansson
Centre for Viking and Medieval Studies, University of Oslo, Norway

Abstract
This article has two aims. First I argue that the old habits of the printed book often have influence on our way of thinking in our work with electronic texts. I stress here the importance of evaluating the new possibilities that the new technological tools provide for philological work. In the second part I present a project that is currently working with the Old Swedish material from the Vadstena monastery in Sweden. Central to this project is to provide electronically encoded transcriptions of the primary sources from Vadstena.

1 Introduction
As the technological tools for producing advanced hypertext editions have developed rapidly in recent years, the possibilities have become obvious to editors of medieval texts. Often, however, philologists end up in endless discussions about software and technicalities, rather than defining areas of research that could be initiated by these new tools. And all too often the focus is on what we see on the computer screen: images of manuscript pages that are not available for any type of search engines, or plain texts that are often only searchable with the simplest computing tools.

The aim of this article is twofold. First I discuss some aspects of electronic texts that still need attention. The old habits of the book in the shape of a manuscript or a printed text still seem to have strong influence on our way of thinking of texts and editions. We therefore, in my opinion, focus too much on the aspects of a text that are traditionally presented in printed editions. As a consequence of this reasoning I go on to outline briefly the relationship between the electronic transcription of primary sources and the edition, printed or electronic, that I think we ought to focus on more in our work. I argue here that all primary sources, i.e. all witnesses of e.g. an Old Norse text, are relevant to researchers in...
different disciplines, and therefore should be made available when we
work with electronic tools.

The second aim is to present two Swedish projects that are relevant in
this context. The purpose of the two projects is twofold: we intend to
make the text from all extant Old Swedish manuscripts from the Vad-
stena monastery available in electronic form within the framework of the
Medieval Nordic Text Archive (Menota; see Haugen in this volume) at the
same time as we shed new light on the production and re-production of
Old Swedish texts within the monastery.

2 Manuscripts, Printed Editions, and Old Habits

The tradition of printed philological editions is several hundred years old.
It is a part of the overall printing tradition that continued the traditions
of the handwritten manuscript. Our way of thinking of book, first page,
pages, indices and so on, is naturally deeply rooted in this tradition. In
philological work we also have a long tradition of how to handle variants
of a work or text. For the editor of medieval texts the task has traditionally
been to analyse the extant text witnesses of a text tradition and recon-
struct the relation between them. From this analysis it has been the aim to
establish a text as close to the hypothetical archetype as possible. In Old
Norse philology, scholars have often used the concept of a best text, i.e.
the extant version that is considered to be closest to the readings of the
archetype has formed the basis of the edition. Variant readings from other
versions have been presented in a critical apparatus below the edited text.
The edited best text has only been corrected when the editor has con-
sidered a reading as corrupt or when text is missing in the best text, and
information about the operation has been added in a footnote. In this
way, the edited text has to a large extent represented an individual
primary source, and users of the edition have been given the opportunity
to trace changes in the edited text in relation to this source. The versions
of the text that are represented only in variant readings, however, are
harder to retrieve from the edition; substantives, i.e. variation that is con-
sidered relevant for the understanding of the text, are often consistently
given in the apparatus, but variants considered as accidentals, i.e. variants
that are seen as e.g. scribal normalization and therefore not relevant for
the understanding of the text, have been silently overlooked.

The limits set by the printed book can be defined primarily in aspects
of space, selection, and time. These three aspects often connect to each
other so that it is difficult to tell one from the other. A common factor for
all three is economy. A printed edition could of course in an ideal world
use as much space as is needed to print all text witnesses, organize the
material in various ways, and use the time needed to do this, but it would
be too expensive, and it would in many cases probably not be worth the
effort. The limits set by space, selection, and time are obvious. In the
preparation of a printed edition these limits are always central to the
editor, who will have to decide on how to choose a best text, how to
collect variants and arrange them on the printed page, and not least, have
Computing Medieval Primary Sources from the Vadstena Monastery

to choose what is not to be presented in the printed edition. The traditional edition will, however, objective our means to analyse versions of a medieval text are, never give the full story, something that can be highly frustrating to philologists who realize that printed editions often do not contain all the relevant information.

3 Computers, Editions, and Electronic Archives

In the printed book, the space given for variants and presentation of different possible solutions is indeed limited. And here we find some important differences between the printed book and the electronic text (or edition). The aspects of space and difficulties of printing are not relevant for electronic editing in the same way, and the ways of presenting texts and variants in the electronic medium set totally different limits. It is also important to stress that a presentation of an electronic edition in a similar way as in the printed edition is in many ways a waste of time and money. In this section I will instead argue for the transcription of primary sources and the building of electronic archives for medieval primary sources that can be used for philological, linguistic, and literary research.

Espen Ore mentions two stages in the editorial process when we are using electronic tools. He differentiates between what he calls ‘teksttilretteleggelse/arkivoppbygging’, i.e. establishing the texts/building of a text archive, and ‘utgaveproduksjon’, i.e. editorial work (Ore, 1999, p. 143). Most editors today have realized that editing the text is easier on the computer than working with excerpts and type-written transcriptions of the text. The first instance, i.e. establishing the text and building a text archive, is perhaps not as obvious. The idea of machine-readable texts and archives with manuscript texts available in electronic form is still a novelty to many philologists, and there are still a lot of preconceptions and misunderstandings. Many philologists still automatically object to someone else using ‘their’ text before it has been edited and printed. The manuscript text is in their opinion only to be used as material for the editor, and they are not easily persuaded to let the text be used for any other purpose. There is still a feeling among many philologists working with Old Norse material that the cumbersome and time-consuming work they have put into their transcriptions and excerpts is not to be used by others. The main reason for this attitude towards new possibilities introduced with the use of computers is, in my opinion, to be found in the old habits discussed above. There is, however, also reason to stress the difference between producing encoded transcriptions of primary sources and establishing an edition. The editorial work is not replaced by the electronic archive, but rather the archive provides a complement to the edition. Often this difference is not acknowledged, and the primary source text and the electronic or printed edition are treated as equal products, which causes some of the reluctance described here.

In traditional textual criticism, the editor works through all manuscripts that contain versions of the text. In some traditions a best text is
chosen, and in others the editor reconstructs an archetypal text from the
text witnesses. In this work, the editor has to work through all the text
witnesses and gather excerpts for the variant apparatus—a tedious but
important work. Instead of following this tradition, however, today all
the text witnesses to a text could be transcribed in full and form the base
for the edition. The material could then be ordered, and normalizations,
editorial emendations, and other editorial arrangements could be made
in accordance with the principles of the tradition; and at the same time
the original versions could be made available in a text archive and used
for other purposes, e.g. for linguistic analysis. In the long run, this would
save us all a lot of work and it would give us material that met linguistic
interests better than an edition where all the decisions are already made
by the editor. The electronic transcription of the primary source is then
only a part of the process to create an edition; the relevant analysis of the
text witnesses is still the work of the trained philologist, while the trans-
scription can be seen as a craft that forms a part of the training for
students of philology.

In this context, the limits set by old habits and economy are often
at play in the discussion. It must, of course, take a lot of time and money
to establish full transcriptions of every individual version of a medieval
text. Here it is advisable to decide from case to case what may be feasible.
In some cases it could be appropriate to immediately transcribe all
versions in full, as for the Old Norse Eddic poems, where there are few
versions, and the manuscript context of the individual version could be
of interest. In the case of the Icelandic medieval law Jónsbók, however,
where there is a large number of versions, it is perhaps only relevant at the
first stage to transcribe the versions of the text that are considered the
main witnesses to the text tradition. In a thorough study of the Jónsbók
tradition, however, e.g. in relation to linguistic development or manu-
script production, it could be argued that other manuscripts should be
added to the corpus. Although, generally, the transcription of the whole
primary source is no more time consuming than making excerpts for the
apparatus.

The problem with space is not really relevant in the electronic format,
and the selection of text and variants could be decisively more complex
than in a printed text. The relation between different versions could be
presented in numerous ways. Editors could give different opportunities
and argue for one, but could also allow for re-analyses made from the text
witnesses (see e.g. Robinson, 2000). In his discussion of different ways of
handling various versions of a text in the electronic edition, Peter Robin-
son mentions the Canterbury Tales edition which presents ‘full trans-
cripts and images of every witness, the best, the worst, the indifferent—all
of them’ (Robinson, 2000, p. 11). He concludes that this presentation of
the witnesses of course is in danger of being incoherent and confusing,
and argues for the establishing of a somewhat modified best text, which
should help the reader of the edition to find a way through all the vari-
ants. I do agree with Peter Robinson that we still need to establish editions
of the various text traditions of the Middle Ages, both in electronic form

96  Literary and Linguistic Computing, Vol. 19, No. 1, 2004
and in printed form, and I think it is relevant to establish principles for how the electronic texts should be presented in an electronic edition. It is, however, important also to stress the primary source in itself as an object of research, something I think Peter Robinson would be the first to agree on. The primary sources presented by the *Canterbury Tales* project are good examples of how the material can be made available for a wide range of disciplines that are not primarily interested in the elusive archetype, but rather focus on the tradition or the individual text witness as a source to a specific context. Linguists, but also historians and literary scholars, could in many cases make more use of the information found in the primary source, which is often impossible to retrieve from a printed edition.

The printed book is not only limited spatially, it is also limited in that it is static. As soon as you have printed the book it will stay the same forever (unless you start to write comments and corrections in the margins, but that is of course a limited form of afterlife). Espen Ore presents the difference between the printed book and the electronic text archive as follows:

> Og mens en papirutgave er død i det øyeblikket hvor den er ferdig, så kan et elektronisk arkiv vedlikeholdes, korrigeres, tilføyes nytt materiale og generelt forbedres (Ore, 1999, p. 144)\(^1\).

The processes involved in the two different kinds of work are, as I have tried to outline in the above, in many cases coincidental. Often the focus is on the editorial work, while the establishing of primary sources for use in an archive is considered worthwhile only as a side product of the editorial work. In the following, however, I will concentrate on the transcription and encoding of primary sources and on the establishing of a text archive.

When we establish electronic texts for the electronic archive there are three main aspects that should be considered, all of which have been treated implicitly in relation to editorial work in the above. It is of course important to stress the need of **accessibility**, i.e. it should be easy for the users to get access to information and material that have so far only been available in archives and libraries. This type of accessibility could be provided with digitized images of manuscript or book pages, which to some extent would take care also of the second aspect, i.e. the **presentation** of the texts on the screen. If a digitized image is available, this admittedly limits the need to use the original. Anyone who just needs to read the manuscript text with accurate pagination will be satisfied with this solution. However the last requirement of electronic texts—**searchability**—is not as easily handled. Problems of consistency immediately present themselves when you start encoding the transcribed texts, and there are many aspects of presentation and accessibility of the searchable texts that need to be solved. It is obvious that when we make an electronically searchable text available, even at the simplest stage, we provide researchers with a valuable tool. And new layers of encoding can further the use of the text.

---

\(^1\) And while a printed edition is dead the same moment it is ready, an electronic edition can be updated, corrected, have new material added, and generally be improved (my translation).
4 Primary Sources from the Vadstena Monastery

In 2001, the Medieval Nordic Text Archive (Menota) was initiated at a meeting in Oslo by representatives from archives and editorial organizations from the Nordic countries. The aim of this organization is to establish an electronic archive for Nordic medieval primary sources in the Vernaculars and in Latin. The principles for text transcription and encoding are outlined by Odd Einar Haugen in this volume. In the following, I present two Swedish projects that over the next four years will work with the Old Swedish manuscript material from the Brigittine monastery in Vadstena, and some of the possibilities of encoded texts that will be explored within these projects. Both projects are funded by The Bank of Sweden Tercentenary Foundation.

The Vadstena Monastery was consecrated in 1384. It was a novelty in the Catholic world in that it was a twin convent with a male and a female contingent. During the following two centuries this monastery became the centre not only of the Brigittine order which spread all over Northern Europe, but it also formed an important cultural centre in the young Swedish state. The end of the Middle Ages, the monastery probably had one of the most impressive book collections in Northern Europe.

The two contingents of the monastery soon established scriptoria that produced a significant number of manuscripts in Old Swedish and Latin. The monks of Vadstena also showed themselves to be important authors, compilers, and translators. The extant manuscripts in Old Swedish comprise some five thousand pages. Their importance for our understanding of the establishing of Swedish as a literate language is obvious. After the Reformation in 1526, and until the closing of the monastery in 1595, it played an insignificant role in Swedish cultural and political life. Already during this period manuscripts were taken away, and in 1595, the library was finally divided and in many cases the manuscripts were destroyed. There is, however, still a significant number of manuscripts in Old Swedish and Latin extant, mainly in Swedish libraries and archives. In the projects described here, we will concentrate on the approximately sixty manuscripts that represent the Old Swedish production at the Vadstena monastery from the period 1384–1540.

To make this material available in electronic form, the work has been divided into two sister projects. In an auxiliary project, The manuscript texts from Vadstena monastery in digitized form, which is based at the University of Oslo in Norway, the focus will be on transcription and encoding of manuscript text. The texts will be transcribed on facsimile level and represented primarily on two levels with the word as the defined basic unit, marked up with the element <w>, the facsimile level <facs> that represents the text as it is read in the manuscript, and the diplomatic level <dipl> that presents the text with expanded abbreviations and minor editorial corrections made in the transcription. A single word, mænniskio ‘man’, will basically be transcribed and encoded as in the following example:

2 For a recent discussion of the Vadstena monastery as a centre of political and ideological importance, see Louise Berglund (2003).
3 The development of this collection, and the means of increasing it, has been studied recently by Anna Fredriksson (1997).
4 For a more thorough presentation of these levels of encoding, see Odd Einar Haugen in this volume.
In this project, tools will also be developed for lemmatization and for morpho-syntactic analysis that will enable us to study variation in morphology and syntax within the Vadstena texts. The \texttt{\textless w\textgreater} element of the above example will be given attributes as to \textit{lemma} or \textit{lexical form} and morphological information as in the following:\footnote{5}

\begin{verbatim}
\texttt{\textless w lemma="menniskia" pos="NFASI">}
\end{verbatim}

The lemma attribute will provide us with a tool to handle both graphic and morphological variation for the token \textit{menniskia}, i.e. we can easily get an overview of all graphic and morphologic forms of a manuscript or a group of manuscripts, while the attribute \texttt{pos} ‘part of speech’ gives us morphological information about the token. Here we get the information that the form is a noun (N) feminine (F) in the accusative (A) singular (S) and indefinite (I).

The material that is established within the project will be available for further quantitative and qualitative analysis not only to scholars engaged in the Vadstena projects, but it will also be made available to other scholars through the \textit{Medieval Nordic Text Archive} mentioned above.

In a parallel research project, \textit{The production of texts and manuscripts in the Vadstena monastery – production, tradition and reception}, the development and use of script in the production and re-production of text within the Vadstena monastery will be studied. This project involves scholars from the University of Umeå and the University of Stockholm in Sweden and the University of Oslo. The theoretical framework is primarily found in New Philology and recent research into the development of literacy.\footnote{6} Therefore the interest is primarily focused on the extant manuscripts and the material evidence they present for processes in the use of script and the book. We concentrate on three areas of interest in this context—\textit{production}, \textit{tradition}, and \textit{reception}. In the following, three studies are outlined to illustrate the use of electronic transcriptions of primary sources within the project.

In New Philology the importance of understanding the manuscript as a means of communication has been stressed (see e.g. Nichols, 1990, 1997). Here it is relevant to study closely the physical manuscript production as well as the text production and re-production in the two scriptoria of Vadstena. The production of manuscripts involves of course a number of craftsmen preparing the parchment, and binding it into books when scribes and miniature painters had done their work. We will concentrate here, however, on the work of translators, authors,
K. G. Johansson

compilers, and scribes, but also illuminations and contemporary bindings will be of interest.

The production and re-production of text in the medieval scriptoria of Europe was a vast undertaking, and in many monasteries it seems to have played an important role in daily life. Authors, translators, and compilers produced original texts, translations into the Vernacular of primarily Latin texts, and compilations and collections of both. This production can in some cases be traced from a first translation through layers of compilations to an end product that shows traces of all stages in this transition. We will encode the texts in ways that enable us to trace the work of compilers and scribes as the text is transformed. Here the results from earlier research on intertextuality and the relation between individual manuscripts can of course be used to establish the relations between various texts and between various stages in text transition for what traditionally has been seen as an individual text. This information will be seen in relation to information retrieved from our own research. On the one hand, the use of original and translated texts in collections and compilations can be mapped and analysed, and we can establish a network of connections between the extant texts over the whole period of time; and on the other hand, profiles for individual compilers and scribes can be established. By combining encoded information about known authors, translators, and compilers with data about scribes, texts, and manuscripts within the corpus, we will be able to shed new light on the use of script and text in the two Vadstena scriptoria.

To be able to study the scribes from Vadstena for the relevant period, we need to carefully define individual scribal hands and distinguish between them in the encoding of the manuscript text. This information will be made searchable in the corpus built by the projects. Here we will encode not only the scribes involved in the original production of the manuscripts, but also later hands that have added to the text; written comments or everyday scribbles will be marked up (see below). For every manuscript, a list of involved hands can be established in the header. For scribes involved in the original production we can define for example a main hand, i.e. a scribe who is responsible for a greater part of a manuscript. For unidentified hands we can order the information as e.g. hand 1, hand 2, and so on. When the corpus of manuscript texts from Vadstena is further developed, each scribe, identified or unidentified, will be assigned a unique number. This will enable us to retrieve all the parts of the corpus where an individual scribe has been at work with a relatively high degree of certainty. From this we can establish a tentative relative chronology for each scribe in relation to datings of the manuscripts where he or she has been active. Information about palaeographic and orthographic use for the individual scribe can be added; this will enable us to refine our means of assessing a relative chronology between scribes and within the production of individual scribes.

In two studies, the relations between translators, compilers, and scribes in a group of manuscripts dated to the second quarter of the 15th century that have earlier been of great interest among Swedish philologists, will

7 The defining of individual hands in the medieval manuscripts has long been a question of contention among Old Norse philologists (see e.g. van Arkel, 1979; Haugen, 1988). The encoding of this information must therefore always be very cautious about the relativity of information about individual scribes.
be investigated in detail with the electronic transcription of the primary sources as a base. The manuscript A 49 in the Royal Library in Stockholm contains, among other things, a translation of the so-called *Baarlams saga ok Josaphats*, a text that was widely known in Scandinavia in the late Middle Ages. The saga is not only available in a fragment of a transcription in the manuscript, but is also preserved *in extenso* in what has been called its concept form, i.e. the translator’s own first version with changes inserted above the line and in the margins. This provides material that is of great importance for our understanding of the relation between the translator’s work and the transcription of a ‘final’ redaction of the text. In a Masters thesis, Maria Arvidsson at the Centre for Viking and Medieval Studies at the University of Oslo, intends to study the work of the translator in relation to the scribes that produced the transcription of the same manuscript. The encoding of every detail in corrections and changes made by the translator can be related to the same instance in the transcription, and thereby variation on all levels can be consistently handled in the machine-readable format. The paleographical and orthographical relation between the translator and the scribes can be mapped as can differences in vocabulary, morphology, and syntax.

In relation to the manuscript A 49, there is a group of texts and manuscripts where scholars have put forward varying views on the translators, compilers, and scribes involved in the production. There has been no disagreement about the dating of texts and manuscripts, but the relation between various texts contained in the manuscripts and between individual monks involved in the production has been a matter of contention. In a Ph.D. thesis, Nils Dverstorp of the Centre for Viking and Medieval Studies at the University of Oslo will provide a study of the scribal hands involved in the production. The study will be based on detailed analysis and encoding of paleographic and orthographic variation of extensive manuscript material with focus on the manuscript A 108 in the Royal Library in Stockholm. From this encoding, a profile can be provided for each individual scribe which will enable Dverstorp to map the scribes in this group and thereby shed new light on the production of the studied manuscripts.

The reception of texts and manuscripts from the Vadstena scriptoria is of course difficult to study. Indications of how the individual manuscript has been used within the monastery by monks or nuns, or outside the monastery walls are of great interest. Manuscript features such as drawings in the margin, damage, and written marginalia that illuminates the use of the manuscript will therefore be carefully encoded. Important information about the use of a manuscript can be gained from marginal notes and scribbles. In earlier editions of medieval manuscripts, some of this information has been presented sporadically in the introductions, but then only marginal notes that were considered to be of relevance to the edited text. In most cases the information was seen as irrelevant. Some scholars, however, found an interest in these notes and scribbles; and in many cases, new information about a manuscript’s provenance was established from them. It has always, however, been difficult to
combine this information from many manuscripts as it is not consistently presented, and in many cases knowledge of this kind of relationship has been in the individual scholar’s memory rather than retrievable for other scholars. Marginalia can be handled within the confines given by the principles for encoding on the word level <w> as <add>, and related to the overall structure of the manuscript related to page <pb n=6r/> and line <lb n=3/>, i.e. the word is situated on the third line of folio 6 recto. In this way we can establish an extensive overview of identified and unidentified scribes in the marginal notes. By ordering them over time (when possible) and in relation to text and/or manuscript, we will be able to access material that is often only briefly presented in printed editions.

The use and reception of a type of text from the 15th century—the prayer book—can be studied in a large number of manuscripts. These were mainly produced by the sisters and therefore give insights into the scriptorium of the nuns in Vadstena. In a Ph.D. thesis from the Centre for Viking and Medieval Studies at the University of Oslo, Ingela Hedström will present a study of the production of prayer books and the users these products were intended for. The large number of manuscripts provides material that can be analysed in both quantitative and qualitative studies. Here the encoding of individual scribal hands can be mapped in relation to a large group of scribes, which will enable Hedström to order the scribes into a relative chronology that can be studied in relation to changes in the genre over time. This study will shed light on the genre of prayer books and their use in the period 1425–1540. It will also provide new and consistent information about the use of script in the convent of nuns in Vadstena.

When the various types of information described for these admittedly only briefly outlined studies are consistently given for all manuscripts within the archive, it will enable us to establish quantitative analyses of the use of abbreviations and of various types of graphs on the paleographic level and relate these to orthographic data. We can further study morphological variation within the established material as well as lexicographic variation. On the material manuscript level we can connect data from various manuscripts and find common phenomena that often escape the scholar. Marginalia that was previously seen as irrelevant can shed new light on the usage of individual manuscripts or groups of manuscripts, and we can get a better understanding of the reception of the manuscripts from the influential Vadstena monastery.

5 Conclusions

When the printed edition was the only available tool to provide a text, the demand for economy was understandable. Editions were reduced to bringing the best text where the variants (and there should not be too many of them) were expelled to the apparatus. The variation that was accepted for the apparatus therefore was limited to what was considered substantial variants, i.e. variants that were seen as relevant for the understanding of the text.
In this article, I have tried to exemplify how the electronic text provides new means of searching and ordering the manuscript material. In the debate on the use of computers in Old Norse philology you sometimes hear the argument that if you put digitized images of the manuscripts on the web you do not need to be so careful in your transcription of the text and your rendering of individual characters or abbreviation signs. In my opinion it is the other way around. Today it is more important than ever to provide the academic world with exact transcriptions of primary source texts, and to ensure that these texts are stringently encoded to facilitate the use of computers for searching large corpora. It is only when the information we are interested in is made available in an electronic format that it can be obtained and ordered, and thereby yield new information. If, for example, information on abbreviation signs is not provided in the encoded text it is as useful, or useless, as the digitized facsimile for the researcher trying to collect quantitative data: no information can be obtained. This also goes for the actual text. It is only when the variation in the manuscript text is rendered and related to a searchable normalization that the electronic text becomes really useful.

It should be stressed, however, that the electronic encoding of texts not only gives us possibilities to organize and search through large amounts of text, it also gives us the option to create a reader’s edition from the encoded information string. It is therefore no longer economically sound to create electronic or printed editions based on the tradition of a best text with variants from other versions in the apparatus. An electronic edition based on primary source transcriptions could provide us with the best of both worlds: an edited text that leads the reader into the text tradition and a complete, searchable database for this tradition.

For the Vadstena projects, we will provide electronic texts according to the lines discussed in this article. In this way we will have taken a step toward establishing a text archive that can be used by scholars from all disciplines. In this short presentation I have given only a few examples of how the establishing of a text archive might change our ways of working with the manuscript material from the Middle Ages, and perhaps more important, might influence our views of the medieval manuscript culture. The two Vadstena projects are planned to develop new knowledge of the monastery of Vadstena, but they will also contribute to the establishing of the Medieval Nordic Text Archive and the development of electronic tools.

References


