Sound, Ink, Bytes: Geographical Information through the Centuries

Øyvind Eide
King's College London

Abstract

This article is based on research carried out as part of the author's PhD studies at the Department of Digital Humanities, King's College, London, and is concerned with the processes of creating and changing texts through different mediums. This article takes the Schnitler protocols, a study commissioned by the Danish government on Norwegian-Swedish/Finish border relations, as an example through which to examine the relationship between different documents, oral, written, print, and digital, and the stylistic and content changes that were likely introduced through each of the transformations from one medium to another.

Keywords / *mots-clés*

Major Peter Schnitler; Media history; Manuscript studies; Textual studies; Textuality /
Major Peter Schnitler; Histoire de medias; Études de manuscript; Études textuelles; Textuality

Øyvind Eide is a Senior Analyst at the Unit for Digital Documentation, University of Oslo, Postboks, 1123 Blindern, 0317 Oslo. Email: oyvind.eide@iln.uio.no

The INKE Research Group comprises over 35 researchers (and their research assistants and postdoctoral fellows) at more than 20 universities in Canada, England, the United States, and Ireland, and across 20 partners in the public and private sectors. INKE is a large-scale, long-term, interdisciplinary project to study the future of books and reading, supported by the Social Sciences and Humanities Research Council of Canada, as well as contributions from participating universities and partners, and bringing together activities associated with book history and textual scholarship, user experience studies, interface design, and prototyping of digital reading environments.

Introduction

Why is the text I am reading exactly as it is? Many text-oriented areas of research, as well as research using texts as source material, ask this question either implicitly or explicitly. It is connected with the question of how a text is created, which is important to examine in order to understand how a source may be interpreted in any historically-oriented study. In this article, a specific text is examined with this question in mind. The research presented here is part of my PhD project at the Centre for Computing in the Humanities, King's College London, and is funded by the Norwegian Research Council.

In the 1740s, Major Peter Schnitler was appointed by the Danish government to explore the border area between the middle and northern parts of Norway, Sweden, and Finland. The manuscript that he handed over to the Danish government in large part consisted of transcripts of local court interviews carried out by Schnitler in order to gather information about the local population and their views on the border areas. The material includes information directly relevant to the border question, as well as general information about the areas in question. The text corresponds to similar materials produced in Europe at the time (Burke, 2000).

This article will analyze Schnitler's material, taking two different developments into consideration: first, the media history of the collection, or, put differently, how the text has been reconfigured into new media through writing, printing, and digitization; and second, the process of creating the text itself, that is, how Schnitler used his collected source material to write comprehensive descriptions of the area in question.

From court interview to digital text

This first section will examine the relationship between the different primary documents – oral, written, printed, and digital – to allow for a better understanding of the stylistic and content changes that were likely introduced through each of the transformations from one medium to another.

Oral to manuscript

Though direct access to the court interviews is impossible, as the sound of the words disappeared the moment they were spoken, there remains written evidence of these interviews. The parts of the original manuscript drawn directly from the court interviews were written during the actual interviews. A simplified court system was used, with only two jury members providing veracity, through signature, of the materials, called protocols. Since it was the law, and Schnitler's group was constantly travelling, the protocols were signed at the time of the interviews, as the jury members stayed behind. According to historical research into court protocols,

such legal documents provide the most accurate record of the actual speech of the witnesses, especially common people (Stretton, 1997; Sandmo, 1992). Though these sources are among the best available, it is likely that many changes were made from the oral interview to its written account. When the witness spoke Sámi, a missionary translated the interview from Sámi to Norwegian, creating new oral texts, since only the translations were recorded. Given that an interview is a dialogue, possible discussions with the witnesses may be included but not explicitly described as such in the protocols, resulting in disparities between the actual interview and its recorded version. Other possible sources of disparities between the oral and written accounts could include people speaking unclearly, or using words or expressions not understood by the scribe, leading to the need for clarifications. It is also possible that witnesses were openly or covertly asked to change their statement. Though there is no evidence of such occurrences, if they did occur, there would be no record of these in the transcripts.

In the protocols, there would be no additions to the content of the stories told by the witnesses, other than in the case of misunderstandings. Anything not said by the witness would be later articulated and only added to the protocol as a statement if the witness confirmed that articulation. In addition, parts of the oral texts are likely to have been removed; for example, if they were seen as irrelevant. Similarly, the form of the statements, that is, the way things were said, is likely to have been changed quite a lot, especially in the translated statements, and will be discussed further in the next section.

In addition to disparities between the oral interviews and their written accounts, it is also possible that other authors, such as colleagues of Schnitler, also contributed to the text, possibly through discussions as he was writing the text. Moreover, the sections in which Schnitler analyzes the sources in aggregate are not based directly on any oral text, and so it can be assumed that he created the text as he was writing it, and it is therefore his own interpretation rather than a direct account of primary sources.

Manuscript to printed text

After Norway left the union with Sweden in 1905, several years of negotiation were needed in order to clear out all the outstanding issues in the relationship between the two now-sovereign states. The question of Sámi reindeer herders moving their flocks across the border proved among the most difficult to settle. Several committees were dedicated to this, among them the Reindeer Grazing Committee of 1907, whose task was to collect and preserve old documents. Two scholars, Just K. Qvigstad from Norway and Karl B. Wiklund from Sweden, led this work, part of which resulted in the publication of just over 400 pages based on Schnitler's manuscript in the National Archive in Oslo (Renbeteskommissionen af 1907, 1909). This publication covered only the latter part of Schnitler's manuscript – the selection process for publication was based on what was needed for border negotiations following the dissolution of the union – and the negotiations only concerned some, not all, parts of the border. The publication of the first part of the manuscript was delayed and only finished in 1961, while a third volume was published in 1982, containing an auxiliary document to the protocols (Schnitler, 1929–1985).

These publications are a faithful reproduction of the manuscript, where the text is neither normalized nor modernized. Though the complex relationship between a manuscript and a transcription will not be dealt with in any detail here, put very

Eide, Øyvind. (2012). Sound, Ink, Bytes: Geographical Information through the Centuries. *Scholarly and Research Communication, 3*(3): 030130, 7 pp.

3

simply, the process of transcription attempts to preserve the text as a string of symbols, with the structure needed to understand it; for example, punctuation. In order to do this, all letters and their capitalization are preserved as in the original manuscript, but the edition is not critical, since only one reading, presumably the most probable one as the transcriber sees it, is recorded. In the case of the Schnitler manuscript, the maps following the manuscript were omitted, as they were deemed too expensive to include, but several helpful tools were added to the text, including a table of contents, page headers, as well as indexes identifying places and persons.

Printed to digital text

The people working in the border commissions immediately saw the importance of the material they collected. In a letter to Emil N. Setälä, Just K. Qvigstad wrote:

> The examinations now being undertaken [...] connected to the reindeer herding case will, when they some day will be available to the public, provide an unusually rich material about the reindeer herding and the living conditions of the Laps. Investigations this minute would never have been undertaken otherwise.[1] [Trans. from Norwegian, my translation]

When a digitization effort was started at the library of Tromsø Museum in 1995 as part of the national Norwegian Documentation Project, the work was in close cooperation with the department of Sámi Ethnography.[2] When material for digitization was chosen, the Schnitler material was considered very important and one volume was finished during the project. The project placed considerable importance on producing digital editions of high quality as well as quantity. Thus, detailed SGML encoding was chosen, but the editions were based on printed versions rather than the original handwritten manuscripts. Thus, the digital version is a new version of the printed books, not a new edition following them.

The aim of the digital version was to reproduce the printed text as minutely as possible. The book pages were scanned and OCR read, followed by several rounds of proofreading. In order to store the structure of the document and add information based on an interpretation of the text, SGML encoding was added to the document. An example of this is that letters set in italic in the printed text were encoded as italic using SGML elements in the digital text (Eide and Sveum, 1998). Eventually, the document was converted from SGML to XML, and the element structure was converted to TEI.
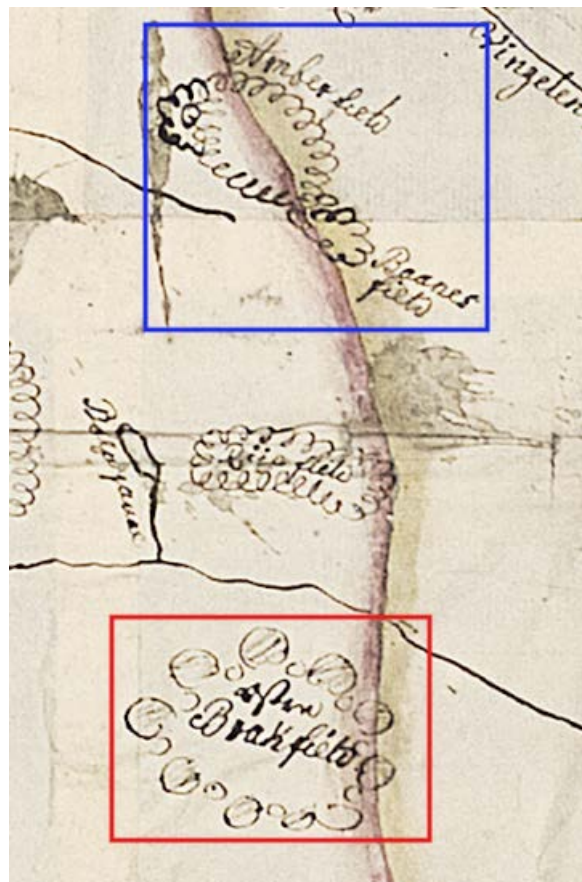
The most important tool that was added in the digital version follows from the medium: an ability to search the text, in free text as well as based on the SGML structure. Based on the SGML version of the text, a software tool was also written in 2002–2003 in order to assist with analysis of the text. The system is no longer used, but the information added to the text through use of the system was exported and is now available for use in connection with the TEI version of the text. The investigations described in the next section were done using this system (Eide, 2004).

From court interview to map

The texts in the Schnitler protocols themselves also have an internal history of information aggregation, performed by Schnitler and his assistants. This internal history consists of the following main steps:

1. Data collection. The court interviews were written down, and older written evidence was collected.

2. Aggregation. Based on the interviews, together with other sources of information, including his own observations, Schnitler described larger areas.

3. Maps. Schnitler drew maps of large areas to indicate where the border should be located based on his sources.

In the course of my PhD project, I will create conceptual models that will be used for analysis of geographical elements of the material captured in the Schnitler volumes. Though the study is currently in progress, there are already obvious differences between individual witnesses' testimonies as they are transcribed and included in the volume. For example, the length of individuals' testimonies and the manner in which they express themselves vary quite a lot. In addition, earlier investigations into word frequencies of these materials, using the analysis tool described above, show demonstrable personal stylistic differences. More specialized methods give similar results. One of these is an analysis of the construction "<place name> <word> <place name>," where the frequency of different words in the middle is counted for different people. The results differentiate individual speakers. Also found were some indications of differentiation among groups of speakers, based on ethnic, professional, and social categories. Groupings based in these differences are conjectural, in part explained by the fact that two different manuscript hands, representing two different scribes, are found in the manuscript as well (Eide, 2004).

Eide, Øyvind. (2012). Sound, Ink, Bytes: Geographical Information through the Centuries. *Scholarly and Research Communication, 3*(3): 030130, 7 pp.

Whatever these differences may be, and whether they are a result of individual or group differences in the expressions of the witnesses, of interpretation, or of different hands in the transcription of the court hearings, it was in Schnitler's interest to remove them when he aggregated the data. His project aimed to include the information from the witnesses with a certain amount of stylistic variation on the one hand, while on the other hand making aggregations in which only the hard facts from the witnesses survived. This process is completed in the maps. It is somewhat preliminary to describe in great detail the differences between descriptions in the texts and their expression on the maps, but it is clear that most contradictions and uncertainty are removed from the material on the map. Two examples will be given of this process, both relating to the portion of one of Schnitler's maps shown in Figure 1.[3]

The first example concerns the location "Østre-Brakfield" (depicted as a red square on the map). The description in Schnitler's aggregation is as follows: he first describes the length and width and some other topographical facts, indicates that "the border here could or would be the 6th witness not state, as the acts show, but it seems likely"[4] (Schnitler, 1929–1985, 1:173–74), and then argues for his view. In the manuscript, the witness' statement is described in his/her own words and then aggregated into Schnitler's overall argument, whereas on the map Schnitler's view is shown without discussion.

The second example concerns the locations of Amberfield and Baanesfield (depicted as blue squares on the map). In his aggregation, Schnitler discusses two different views held by groups of witnesses living in different parishes, in which two different mountains are considered to mark the border. Though Schnitler states that he is not in a position to choose between these two views, as he has not been able to gather the two groups of witnesses together to reconcile the matter, he also argues that the most likely solution is to designate Amberfield as the border mountain. Nonetheless, they are both included in the headword in his list of mountains with an "or" between them, and both are included on the map (Schnitler, 1929-1985). Whether their difference in size on the map is reflective of his view of the correct choice is unknown.

Conclusion

The transition from sound to digital form aims to store the informational contents of what was said and the way it was expressed. In the first part of this process in the Schnitler protocols, that is, the transcription of the interviews, there is reason to believe that the scribes knew the limitations of their method, especially when the actual words were translated for them. At the same time, they were most likely able to use expressions such as "recording what the witness said" to describe their work. In the rest of the transformations, from manuscript onwards to digital TEI encoded text, the aim was to store all marks in the previous medium with any relevant information value, and remove the rest. The transcribers had to choose between what was seen as relevant and what was not. One example of what was kept is the use of capitalization in the printed and digital texts based on the usage in the manuscript.

The development from witness transcripts, to aggregation, to map form is done with a different aim. The goal is to use the different sources (the witness transcriptions, the older printed sources, and Schnitler's own observations and knowledge) for an accurate description of a geographic area. I am deliberately not using the word "truthful" here,

because, as a representative of the Danish king, Schnitler was far from neutral. Though we cannot know his thoughts throughout this work, it is likely that he tried to express the reality in the field, but when making judgments based on his aggregated data, as we saw in the last section, he must have taken his own role and loyalty into consideration, whether consciously or not. Despite this, the aim of Schnitler's work was to draw on sources to form a conclusion, not very different from scholarly work. Thus, the first process described is a process of information preservation, while the second is a process of source-based information creation.

Notes

1. National Archives of Finland, Prof. Setälä's private archive. Letter dated "Kristiania 15/10 1911". http://www.dokpro.uio.no/qvigstad/brev.html .

2. Web http://www.dokpro.uio.no

3. Norwegian National Archives, The border archive, Map 120. A facsimile is published on a CD following a publication of another Schnitler manuscript (Mordt, 2008). CD-ROM.

4. Translated from Danish, my translation.

References

Eide, Øyvind, & Sveum, Tor. (1998). *Dokumentasjonsprosjektet ved Universitetsbiblioteket i Tromsø. Rapport.* [The Documentation Project at the University Library of Tromsø. Report.]. Tromsø, Norway.

Eide, Øyvind. (2004). *Fra SGML til begrunnede påstander om verden : et system for analyse av geografiske resonnementer uttrykt i historiske tekster* [From SGML to motivated statements about the world : a system for analysis of geographic reasoning expressed in historical texts]. Oslo, Norway: University of Oslo.

Burke, Peter. (2000). *A social history of knowledge: From Gutenberg to Diderot.* Cambridge, UK: Polity Press.

Mordt, Gerd (Ed.). (2008). *"Møre og Romsdal, Sør-Trøndelag, Nord-Trøndelag, Nordland,Troms." Norge i 1743 : innberetninger som svar på 43 spørsmål fra Danske Kanselli* [Norway, in 1743, reports answering 43 questions from the Danish chancel], (p. 5). Oslo, Norway: Solum.

Renbeteskommisjonen af 1907 [The Reindeer Herding Commission of 1907]. (1909). *Dokumenter angaaende flytlapperne m.m.* [Documents regarding the nomadic laps et cetera.] Kristiania, Norway.

Sandmo, Erling. (1992). *Tingets tenkemåter. Kriminalitet og rettssaker i Rendalen, 1763–97.* [The thinking of the court. Crime and trial in Rendalen, 1763–97]. Oslo, Norway: Tingbokprosjektet.

Schnitler, Peter. (1929-85). *Major Peter Schnitlers grenseeksaminasjonsprotokoller 1742–1745.* [Major Peter Schnitler's Border Examination Protocols 1742–45] Oslo Norway: I Kommisjon Hos A/S Bokcentralen.

Stretton, Tim. (1997). Social historians and the records of litigation. In S. Sogner (Ed.), *Fact, fiction and forensic evidence: The potential of judicial sources for historical research in the early modern period* (pp. 15-34). Oslo, Norway: University of Oslo.

Eide, Øyvind. (2012). Sound, Ink, Bytes: Geographical Information through the Centuries. *Scholarly and Research Communication, 3*(3): 030130, 7 pp.

7