Trawling and trolling for terrorists in the digital Gulf of Bothnia :
Cross-lingual text mining for the emergence of terrorism in
þÿ S w e d i s h   a n d   F i n n i s h   n e w s p a p e r s ,   1 7 8 0   1 9 2 6

Fridlund, Mats

de Gruyter
2022-10-24

Mats Fridlund*, Daniel Brodén, Tommi Jauhiainen,
Leena Malkki, Leif-Jöran Olsson, and Lars Borin

Trawling and Trolling for Terrorists in the Digital Gulf of Bothnia

Cross-lingual Text Mining for the Emergence of Terrorism in Swedish and Finnish Newspapers, 1780–1926

Abstract: In pursuing the historical emergence of the discourse on terrorism, this study trawls the "digital Gulf of Bothnia" in the form of a corpus of combined Swedish and Finnish digitized newspaper texts. Through a cross-lingual exploration of the uses of the concept of terrorism in historical Swedish and Finnish news, we examine meanings anchored in the two culturally close but still decidedly different national political contexts. The study is an outcome of an integrative interdisciplinary effort

---

Note: This chapter uses the following notational conventions for source-language data. Source-language expressions are set in italics accompanied by English glosses in single quotes (also used when a gloss is repeated without the source-language expression): *hirmutyö* 'atrocity'. Furthermore, English expressions in single quotes are used to refer to the central concepts under discussion, for example, 'terrorism'. Some source-language material passages are provided as historical arguments rather than linguistic examples. Translations of such passages – made by the chapter's authors – are rendered in double quotes.

***Corresponding author: Mats Fridlund,** Centre for Digital Humanities, Department of Literature, History of Ideas and Religion, University of Gothenburg, Gothenburg, Sweden,
e-mail: mats.fridlund@lir.gu.se
Daniel Brodén, Centre for Digital Humanities, Department of Literature, History of Ideas and Religion, University of Gothenburg, Gothenburg, Sweden, e-mail: daniel.broden@lir.gu.se
Tommi Jauhiainen, Department of Digital Humanities, University of Helsinki, Helsinki, Finland,
e-mail: tommi.jauhiainen@helsinki.fi
Leena Malkki, Centre for European Studies, University of Helsinki, Helsinki, Finland,
e-mail: leena.malkki@helsinki.fi
Leif-Jöran Olsson, Språkbanken Text, Department of Swedish, Multilingualism,
Language Technology, University of Gothenburg, Gothenburg, Sweden,
e-mail: leif-joran.olsson@svenska.gu.se
Lars Borin, Språkbanken Text, Department of Swedish, Multilingualism, Language Technology,
University of Gothenburg, Gothenburg, Sweden, e-mail: lars.borin@svenska.gu.se

by Swe-Clarin, using resources accessible through the CLARIN infrastructure to enrich scholarship in the humanities. The capabilities of the corpus tool Korp enable us to affirm prior research on the conceptual history of terrorism, but also to suggest a complex and diverse picture of the connotations of terrorism, both as state and sub-state violence up until the 20th century. At the same time, the study allows us to explore the potentials of cross-lingual text mining for historical analysis of national online newspaper corpora provided by Swe-Clarin and FIN-CLARIN.

Keywords: history of terrorism, digital history, Korp, comparative corpus studies

1 Introduction

The development of large-scale digitization initiatives (LSDIs) and language technology (LT) infrastructures has contributed significantly to opening up historical big data for research, allowing scholars to pursue large-scale research questions and explore past phenomena by "trawling" through massive amounts of text. (Weller 2013; Graham, Milligan, and Weingart 2016; Paju, Oiva, and Fridlund 2020). However, critical commentators such as Tahmasebi et al. (2019) argue that many such projects are deficient, being strongly biased either towards data science or humanities, and thus lacking in either technical and linguistic proficiency for utilizing the potential of big data text analysis or appropriate humanistic domain knowledge to evaluate whether the results are pertinent.

The present study is part of an integrative interdisciplinary initiative to overcome such limitations launched by the Swedish CLARIN node (Swe-Clarin: sweclarin.se), which includes pilot projects where researchers in natural language processing collaborate closely with humanities scholars to explore the broad research potential of LT-based e-science tools (see Viklund and Borin 2016; Karsvall and Borin 2018). The chapter builds and expands on two preliminary studies coordinated by Swe-Clarin that used text mining of Swedish-language newspaper corpora from the late 18th to the early 20th century to explore the historical emergence and evolution of terrorism (Fridlund et al. 2019, 2020). This study deepens these earlier efforts through a cross-lingual investigation of Swedish and Finnish newspaper discourse, involving researchers from Sweden and Finland. A key point is the mutually beneficial outcome of the interdisciplinary collaboration: while the terrorism scholars produce a complex historical analysis of the results from the LT resources, the humanistic research questions provide a cross-lingual use case for the data analysts.

This chapter proceeds by discussing the LT tools and the Swedish- and Finnish-language newspaper corpora we have used to approach and trawl through

the historical newspaper discourses on terrorism. Following that, the chapter turns to the analysis of the attributions of terrorism in our dataset. This is the longest section, as it concerns the exploration of a range of attributions of both state terrorism and sub-state terrorism that are explored through a combination of distant and close reading. The distant reading discussion is, to some extent, centered around attributions of nationality, which proved to be a particularly significant factor. After a comprehensive exploration, we turn to a case study using close readings of the emergence of domestic attributions of terrorism in Finland in the early 20th century. Here, we abandon non-selective trawling in favour of directed "trolling", or precision searches, to catch specific uses of the term related to different contexts from those in Russia, which influenced its early uses. We conclude by emphasizing our more significant findings and also make some reflections on the potentials of evaluative historical studies based on cross-lingual text corpora.

2 Computing the history of terrorisms

The first known uses of 'terrorism' as an exclusive description of violent state practices (Erlenbusch 2015) is from the French Revolution's Reign of Terror in 1794. The traditional scholarly view is that the later so-called sub-state terrorism, or "rebel" terrorism, which today is primarily associated with the concept, was first introduced as a tactic by Russian social revolutionaries in the late 1870s (sometimes referred to as "the Russian method") which had already spread to Western Europe, Asia, and America during the 19th century, and reached the rest of the world in the early 20th century (Ker 1917; Law 2009; Miller 2013; Sageman 2017). Notably, this historical picture has essentially been based on close reading of primary and secondary textual sources. Only rarely have researchers examined the emergence of new forms of terrorism through a quantitative approach, including newspaper text mining (cf. Ditrych 2011, 2014; Jensen 2018). However, combining historical, theoretical, and technical expertise, the present study performs both distant reading and close reading analysis of the development of the historical discourse on terrorism in Swedish and Finnish newspapers.

A central component of the research presented in this chapter is *conceptual history*, a field of inquiry which by necessity must grapple with the vexed questions of the nature of concepts and their relationship to the words that express them (Ifversen 2011). As a concrete illustration of this issue, Princeton WordNet (PWN; Fellbaum 1998), a lexical resource for English heavily used in all kinds of text-processing and text-understanding applications, makes a distinction between concepts (called *synsets*), words, and word senses. Almost half – or close to 54,000 –

of the PWN synsets consist of more than one word sense and consequently find expression in more than one way in texts. On average, such senses belong to almost three synsets each. While concepts are not words, scholars of conceptual history have arguably tended to downplay this distinction, investigating the use of particular words as a stand-in for the concepts these words (purportedly) express. A concern that tends to emerge and which is not typically addressed in disciplines that mainly rely on close reading and "thick description", however, is that of typicality or representativeness: how is a posited concept typically expressed – in the way suggested by the researcher's "pre-empirical" intuition or by some other means? In our case, the combination of, on the one hand, distant reading and, on the other, close reading and abduction on the basis of individual instances embedded in rich contexts provides a way of working through this problem.

2.1 Purpose and aims

The main aim of the study is to evaluate the established hypothesis that the modern meaning of terrorism as sub-state political violence did not emerge outside the revolutionary context of Russia until the 20th century (see Fridlund 2018; Jensen 2018). Also included in this aim is the assessment of how the original meaning of state terrorism persisted as an integral historical part of the concept of terrorism.

Sweden and Finland provide a pertinent combination of historical contexts for exploring the development of the discourse on terrorism. Notably, the two countries share a common history – Finland was a part of Sweden from medieval times until 1809 and still retains Swedish as one of its two official languages (about 5% of the Finnish population have Swedish as their mother tongue and a significant number of Finns are bilingual). At the same time, the national conditions have been different when it comes to political violence. Sweden, like many other Western European countries, experienced few instances of terrorism during the period in focus (one bombing and one shooting 1908–1909). However, Finland, following a war in 1809, was incorporated into the Russian Empire as a Grand Duchy (1809–1917), bringing the country closer to the Russian political culture and its revolutionary and terroristic contexts. In fact, Finland suffered a domestic terrorist campaign 1904–1907 in one of the earliest examples of sub-state terrorism (Kujala 1992; Fridlund and Sallamaa 2016; Jensen 2018). In this sense, a dual focus on Sweden and Finland provides deep and different historical horizons on the phenomenon of terrorism.

To trace the development of the discourse on terrorism in Sweden and Finland during the period in focus, our study mainly pursues two research questions: (1) what attributions of terrorism were made in the two countries; (2) to what extent

did the "original" meanings of terrorism persist and other meanings emerge? This encompasses an interest in the historical political events and practices that terrorism has been associated with in the Swedish and Finnish contexts.

3 LT-driven trawling in shared waters

To pursue the historical emergence on terrorism, our study maps the meanings associated with terrorism in Swedish and Finnish newspapers from the late 18th century to the early 20th century. To use a familiar metaphor for text mining in digital humanities, our study "trawls" (see Tangherlini and Leonard 2013) through the vast and deep "digital Gulf of Bothnia" of digitized historical Swedish and Finnish newspapers. This gulf is the Baltic Sea's northernmost part, consisting of Swedish and Finnish territorial waters and a shared body of water in between (which before 1809 was domestic Swedish waters); see Figure 1. Similarly, the body of texts we trawl through for occurrences of words (such as 'terrorism' or 'terrorists') consists of uniquely domestic Swedish and Finnish news as well as shared "transnational" news published in newspapers in both countries. The specific resources we use for trawling are the corpus search tool Korp and historical Swedish- and Finnish-language newspaper corpora provided by the National Swedish Language Bank (Nationella språkbanken) and the Language Bank of Finland (Kielipankki/Språkbanken).
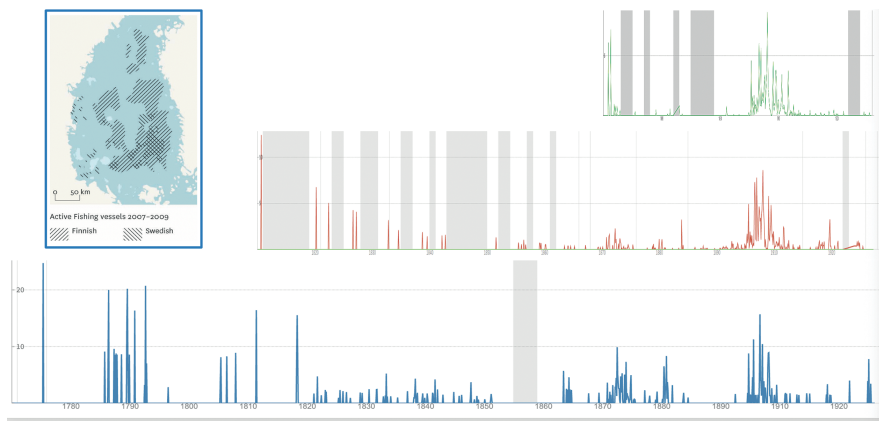

Figure 1: Gulf of Bothnia with partly overlapping Swedish and Finnish fishing waters. From Backer and Frias (2013:52). Courtesy of the Helsinki Commission. Trend graphs showing hits in Swedish and Finnish corpora for *terrorist/terrorism* for seKorp 1780–1926 (bottom), fiKorpSV 1805–1925 (middle) and *terroristi/terrorismi* for fiKorpFI 1880–1925 (top).

3.1 Korp for Swedish and Finnish

Korp (Borin, Forsberg, and Roxendal 2012) is a sophisticated corpus search tool with modular design and an online search interface that, although designed to fulfill the research needs of linguists, has proven useful in addressing humanities research questions.

Its interface allows searches and queries based on automatic linguistic annotations with structured result presentations: a *contextual hit list* or *KWIC* (keyword in context); *statistical data* of keyword occurrences in sub-corpora allowing creation of *trend graphs* plotting relative frequencies over time for text words, lemmas (dictionary headwords), or other linguistic items; a so-called *word picture* presenting statistically prominent fillers of selected syntactic dependency relations of a keyword, for instance typical subjects and objects of a verb, and nominal premodifiers (e.g., adjectives) and post-modifiers (prepositional phrases or main verbs of relative clauses); see Figure 2. The word picture can be used as a topical map to guide users to closer readings of the corpus. Korp also supports navigation between the statistics, trend-graph, and word picture views, and the KWIC view allows close reading of individual hits in their newspaper article context.

The original Swedish version of Korp (from now on seKorp) is developed and maintained by Språkbanken Text (the Swedish Language Bank's Text Division), a national language technology infrastructure development center and the coordinating node of Swe-Clarin, while the Korp implementation in the Language Bank of Finland (fiKorp) is a modification of seKorp by researchers at FIN-CLARIN. Notably, the two Korp configurations are somewhat different. For example, data-wise, fiKorp gives an order of magnitude higher frequencies for some of the terms in its Swedish newspaper subcorpora (fiKorpSV), partly due to better OCR quality (Figure 1). At the same time, feature-wise, direct multi-lemma comparison is not possible for fiKorpSV, although peaks for individual terms in the trend graphs can nevertheless indicate tendencies for further investigation.

3.2 Newspaper corpora in two languages

The Swedish newspaper corpus used for our study, Kubhist, is a large collection of historical newspapers of Sweden from the late 18th to the early 20th century digitized by the National Library of Sweden, containing about 5.5 billion words

(Adesam, Dannélls, and Tahmasebi 2019).[1] While Kubhist is smaller than, for example, the Google Books dataset, it distinguishes itself from many historical newspaper LSDIs such as impresso, Europeana, and NewsEye in being linguistically annotated on several levels (lexical, morphological, lexical-semantic, syntactic, named entities, etc.). The annotation tools also draw on high-quality lexical resources (historical and modern), which compensates for its smaller size, relatively speaking (Borin and Johansson 2014; Tahmasebi et al. 2015; Adesam, Dannélls, and Tahmasebi 2019).

One notable omission in Kubhist, in the context of this study, is that it does not include any newspapers from the Swedish region of Finland. However, the Finnish Newspaper and Periodical Corpus of the National Library of Finland (NLF),[2] includes newspapers of Finland both in Finnish (NLF 2011a) and Swedish (NLF 2011b) from the period chosen. As Finland was a part of the Swedish realm until 1809, Swedish was for a long time the dominant written language in Finland, even during the Russian reign 1809–1917 (although its influence waned during the 20th century). The current version of the corpus includes 5.2 billion words in Finnish and 3.5 billion words in Swedish. Like Kubhist, the Finnish corpus is not complete for any given period of time, as the NLF digitization effort has not been comprehensive.[3]

It should be noted that the period for analysis, 1780–1926, is chosen for both historical and pragmatic reasons with regard to the corpora used. While Kubhist covers the years 1749–1926, the corpus is complete from 1780, that is, about ten years before the French Revolution, which "birthed" the concept of terrorism, providing us with a baseline against which to trace its development. Kubhist also ends in 1926 due to copyright restrictions, effectively limiting the analysis to the period chosen.

---

1 Kubhist also forms one of the major components of the Swedish Diachronic Corpus, a Swe-Clarin initiative described elsewhere in this volume (Pettersson and Borin 2022).
2 http://urn.fi/urn:nbn:fi:lb-201405276
3 For example, the largest Finnish language newspaper *Helsingin Sanomat* is missing issues from 1913 onwards and the Swedish language newspaper *Helsingfors Tidningar* is missing at least 1,218 issues from 1860–1864. However, a significantly expanded version of the corpus, including many of the missing issues, is currently under construction. Preliminary figures indicate that the number of Finnish publications in the corpus will increase by c. 740,000 and the number of Swedish publications by c. 120,000. See http://urn.fi/urn:nbn:fi:lb-202009152, which has a link to the list of new issues in all languages.

4 Reading emergent forms of terrorism

To reach and capture the wide context of terrorism during the period 1780–1926, we formulated Korp queries combining the search terms *terrorist/terroristi* (for 1780–1926 there were 259, 1,364, and 2,629 hits in the Swedish, Finnish-Swedish, and Finnish corpora respectively) and *terrorism/terrorismi* (570, 2,361, and 1,633 hits). Figure 1 (see above) shows the three graphs for *terrorist/terrorism* in seKorp (bottom/blue) 1780–1926 and fiKorpSV (middle/red) 1805–1925 and for *terroristi/terrorismi* in fiKorpFI (top/green) 1880–1925 (there were no hits in the Finnish-Swedish corpus before 1805 or after 1925 and in the Finnish corpus before 1880 and after 1925). In Figure 3, the left column shows the graph for 'terrorist' from bottom to top from seKorp (red) 1780–1926, fiKorpSV (blue) 1805–1925, and fiKorpFI (green) 1880–1925, which have similar profile shapes.

It is important to note the difficulty in generalizing about how common different terrorism/terrorist attributions were, based on the *quantity* of these and other hits, due to the fact that a high number of attributions might refer to multiple descriptions of one specific event in different newspapers. The reuse of near identical news texts in different newspapers (often without attributing the original source) was also a common and accepted practice (Salmi et al. 2013). Thus, what is most relevant in the following is not the quantity of occurrences of a certain attribution, but the qualitative *existence* of such an attribution to terrorism.

One should also note that when 'terrorism' is used in the material to designate certain activities, it is often not clear how violent or lethal these were. As Claudia Verhoeven writes, the verb *terrorise* "suggests the use, power, and violence of the *word*, not the act: 'to terrorize' means to force or provoke certain actions via threats and intimidation, but does not automatically imply physical violence" (Verhoeven 2004: 18). Consequently, it may be difficult to distinguish the type and the "quality" of the terror that the various attributions of 'terrorism' refer to.

4.1 Word picture analysis and terrorism in contexts

The use of Korp's "word picture" function enables a comparison between Swedish and Finnish newspapers, while uneven as it is not activated in the Swedish-language part (fiKorpSV) of the fiKorp NLF newspaper and periodical collection (there are plans to eventually extend this functionality to all fiKorp corpora). Moreover, word pictures are only activated for searches in the collections's Finnish language part when using the "simple search" function, which does not make it possible to distinguish between newspapers and periodicals. The multitude of Finnish-language word forms in word pictures also makes it more difficult

to interpret the results, compared to similar Swedish-language word pictures. For example, the Swedish word for 'Russian', *rysk* (adj),[4] corresponds to more than 50 different words in the Finnish *terroristi* word picture. Many of these words include slight OCR errors, which makes using automated methods difficult, such as the form *venäläiset* in modern Finnish – nominative plural of the adjective *venäläinen* 'Russian' – which is found in the following two erroneous forms on the top 44 list as attributes for the word *terroristi*: *menäläiset* and *roenäläifet*, in addition to the older spellings *wenäläiset* and *Wenäläiset*.

Nevertheless, to compare the Swedish and Finnish contexts, we used the results from the seKorp word pictures and manually scanned the 3,725 concordances for *terrorist* and *terrorism* in the fiKorpSV KWIC view for significant pre- and post-modifiers, including nationalities, locations, and gender. As another form of analysis, we performed close readings in the KWIC view of seKorp and fiKorp, selectively reading the digitized newspaper articles where the hits occurred. While this strategy potentially missed out on some uniquely Finnish uses of terms that Finnish-Swedish word pictures might have revealed, it arguably strengthened the comparative aspect to a feasible level.

Through the word picture function, we were able to examine national or ethnic attributes given to *terrorism* and *terrorist* and to determine that the dominant terrorism-related *national context* in Swedish and Finnish newspapers, as expected, is 'Russian' with 15 hits (*rysk*) in seKorp and some 200 (*venäjän*) in fiKorpFI. Additionally, 25 other terrorist nationalities were identified, although it is difficult to determine how prominent these actually were, due to the limited number of hits.

The seKorp word pictures attributed 8 nationalities (Russian, Chinese, Finnish, French, German, Hungarian, Irish, and Polish) with Chinese as the only unique seKorp attribution (see Figure 2). KWIC readings of fiKorpSV resulted in 23 nationalities (Russian, American, Armenian, Baltic, Bengali, Bulgarian, Czechian, Croatian, Finnish, French, German, Grusinian [Georgian], Hungarian, Indian, Irish, Italian, Latvian, Polish, Prussian, Romanian, Spanish, Vatican, Wallachian) of which 11 were unique (American, Baltic, Bengali, Croatian, Czechian, Grusinian, Latvian, Prussian, Romanian, Vatican, Wallachian) and several were also subjects of the Russian empire. The Finnish-language newspapers attributed 14 nationalities (Russian, Argentinian, Armenian, South African, Bulgarian, Estonian, Greek, Indian, Irish, Livonian [Estonian], Polish, Serbian, Spanish, Turkish) of which 4 were unique (Argentinian, Greek, Serbian, Turkish) and 3 were connected to the Ottoman empire. In total, 27 nationalities were attributed to terrorism in our dataset (17 uniquely attributed in newspapers of one of the three language contexts).

---

4 Also *ryss* (noun), but not in this context.

4.2 Broadening of regime terrorism

The neglect of state or regime terrorism "as a subject for systematic and sustained research" is said to be a "perennial criticism" of terrorism studies (Jackson 2008: 377), and even more so the lack of historical studies (for an exception, see Miller 2013). However, through our data-driven approach, we were able to follow the discursive trajectory of state terrorism and to determine that in both Sweden and Finland terrorism remained strongly associated with political violence and repression perpetrated by regimes at least up until the early 20th century.

Word picture attributive modifiers, such as 'monarchical', 'oligarchic', 'dictatorial', 'military', 'official', 'statist', and 'government/al' (*hallitus/hallinnollinen*) terrorism, clearly show that regimes were held to be agents of terrorism during the period. However, notably, our KWIC readings of these results showed that some attributes, such as 'ministerial' and 'autocratic' terrorism, referred to what may be called "soft" regime terrorism, involving intimidation, harassment, or repression but rarely physical violence. This could be taken as an indication that the meanings of 'terrorism' were widening during the period so that at times they took on metaphorical meanings, as when used in a Swedish political context to describe a statist 'inquisitorial' terrorism or when Finnish temperance advocates were criticized for their allegedly fanatic actions which were labelled 'sobriety terrorism'.

Our analysis of terrorism's national-ethnic attributions shows that the concept of 'terrorism' was used both for *state* regime terrorism and *sub-state* rebel terrorism during the period in focus. For example, the frequent 'Russian' attributions point toward Russian regime terrorism as well as the revolutionary rebel terrorism campaigns of the 1880s and early 1900s.

That state terrorism remained a significant part of the discourse is further indicated by the nations and regions associated with *terrorism* and *terrorist* in Korp's word pictures (for terrorists 'in France' see Figure 2). The findings contain many national forms of state terrorism during the period 1848–1867. The 'German' terrorism found in seKorp refers to activities by a Prussian army in 1848 in the occupied Danish Duchy Schleswig-Holstein. Similarly, the 'Hungarian' and 'Polish' terrorisms are connected to war and occupation following the failed 1849 Hungarian revolution, where an occupying Austrian regime in 1850 was accused of terrorism, as was a temporary rebel regime 1863 in Russian Poland. Such regime terrorism in contested regions, domestic or occupied, accounts for several other regional attributions, such as Armenian, Baltic, Czechian, Finnish, Grusinian, Indian, Latvian, Romanian, Svecoman, and Wallachian terrorism, many of these nationalities belonging to European and Asian empires and the Russian empire in particular, consolidating the significance of the Russian context in the development of the terrorism discourse.
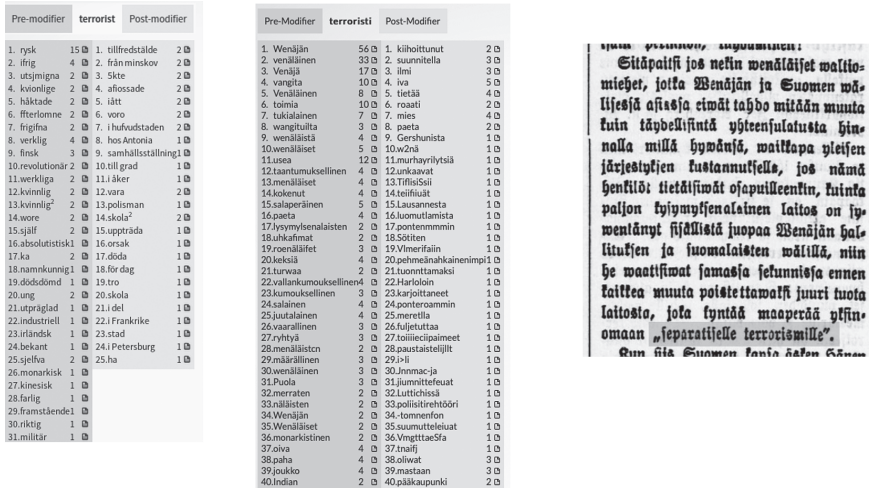
Left word picture — terrorist (seKorp)

Pre-modifier		terrorist	Post-modifier	
1. rysk	15		1. tillfredstälde	2
2. ifrig	4		2. från minskov	2
3. utsjmigna	2		3. 5kte	2
4. kvionlige	2		4. afiossade	2
5. håktade	2		5. iått	2
6. fterlomne	2		6. voro	2
7. frigifna	2		7. i hufvudstaden	2
8. verklig	4		8. hos Antonia	1
9. finsk	3		9. samhällsställning	1
10. revolutionär	2		10. till grad	1
11. werkliga	2		11. i åker	1
12. kvinnlig	2		12. vara	2
13. kvinnlig²	2		13. polisman	1
14. wore	2		14. skola²	1
15. själf	2		15. uppträda	1
16. absolutistisk	1		16. orsak	1
17. ka	2		17. döda	1
18. namnkunnig	1		18. för dag	1
19. dödsdömd	1		19. tro	1
20. ung			20. skola	1
21. utpräglad	1		21. i del	1
22. industriell	1		22. i Frankrike	1
23. irländsk	1		23. stad	1
24. bekant	1		24. i Petersburg	1
25. sjelfva	2		25. ha	1
26. monarkisk	1			
27. kinesisk	1			
28. farlig	1			
29. framstående	1			
30. riktig	1			
31. militär	1			

Right word picture — terroristi (fiKorpFI)

Pre-Modifier		terroristi	Post-Modifier	
1. Wenäjän	56		1. kiihoittunut	2
2. venäläinen	33		2. suunnitella	3
3. Venäjä	17		3. ilmi	3
4. vangita	10		4. iva	5
5. Venäläinen	8		5. tietää	4
6. toimia	10		6. roaati	2
7. tukialaisen	7		7. mies	4
8. wangituilta	3		8. paeta	2
9. wenäläistä	4		9. Gershunista	1
10. wenäläiset	5		10. w2nä	1
11. usea	12		11. murhayrityksiä	1
12. taantumuksellinen	4		12. unkaavat	1
13. menäläisen	4		13. TifliisiSsii	1
14. kokenut	4		14. teifihiät	1
15. salaperäinen	5		15. Lausannesta	1
16. paeta	4		16. luomutlamista	1
17. lysymylsenalaisten	2		17. pontenmmmin	1
18. uhkafimat	2		18. Sötiten	1
19. roenäläifet	3		19. VImerifäin	1
20. keksiä	4		20. pehmeänahkaineninpi	1
21. turwaa	4		21. tuonnttamaksi	1
22. vallankumouksellinen	4		22. Harloloin	1
23. kumouksellinen	3		23. karjoittaneet	1
24. salainen	4		24. ponteroammin	1
25. juutalainen	4		25. meretlla	1
26. vaarallinen	3		26. fuljetuttaa	1
27. ryhtyä	3		27. toiiiieciipaimeet	1
28. menäläistcn	2		28. paustaisteljilt	1
29. määrällinen	3		29. i>li	1
30. wenäläinen	3		30. Jnnmac-ja	1
31. Puola	3		31. jiumnittefeuat	1
32. merraten	2		32. Luttichissä	1
33. nhäläisten	2		33. poliisitirehtööri	1
34. Wenäjän	2		34. -tomnenfon	1
35. Wenäläiset	2		35. suumutteleiuat	1
36. monarkistinen	2		36. VmgtttaeSfa	1
37. oiva	4		37. tnaifj	1
38. paha	4		38. oliwat	1
39. joukko	4		39. mastaan	3
40. Indian	2		40. pääkaupunki	2

Sitäpaitsi jos nelin wenäläiset waltio-
miehet, jotka Wenäjän ja Suomen wä-
lisessä asiassa eiwät tahbo mitään muuta
tuin täybellisintä yhteensulatusta hin-
nalla millä hywänsä, waikkapa yleisen
järjestyjen kustannuffella, jos nämä
henkilöt tietäisiwät osapuilleentin, kuinka
paljon työyympfenalainen laitos on ty-
wentänyt sisällistä juopaa Wenäjän hal-
lituffen ja suomalaisten wälillä, niin
he waatiswat samasfa sefunnissa ennen
kaiffea muuta poistettawafsi juuri tuota
laitosta, joka kyntää maaperää ylfin-
omaan "separatijelle terrorismille".
Kun siis Suomen kania äslen hänen...

Figure 2: Distant and close readings of terrorists and terrorism in context. Word picture in seKorp and fiKorpFI showing pre- and post-modifiers for *terrorist* (left) and *terroristi*. Finnish editorial warning about 'separatist terrorism' in *Vaasa* (12 September 1905). From *digi. kansalliskirjasto.fi*.

4.3 Diversification of rebel terrorism

It took time, however, for the phenomenon of sub-state or rebel terrorism, to acquire a wider meaning beyond the Russian political context. This broadening of the 'terrorism' concept to include national sub-state political militants other than Russians, as well as violent anarchists and anti-imperial revolutionaries, can be studied by comparing and contrasting occurrences of 'terrorists' in our dataset with closely associated terms for sub-state actors traditionally regarded as among the period's prominent practitioners of political violence.

The most prominent such political militants were *anarkist/anarkisti* 'anarchist' (3,028, 20,837, and 22,481 hits), *nihilist/nihilisti* (1,660, 3,113, and 3,148 hits) ('nihilists' and 'nihilism' were up until the 1890s often used as synonyms for Russian social revolutionaries and their ideologies), and *revolutionär/vallankumouksellinen* 'revolutionary' (noun: 1,285, 9,618, and 0 hits) (the fiKorpFI count is zero for 'revolutionary' due to incorrect part of speech tagging in the corpus; all the *vallankumouksellinen* nouns are tagged as adjectives).

Figure 3 shows in its right column a 'terrorist' trend graph and trend graphs for the other political militants for the period 1848–1920, which covers numerous turbulent events, including the European political upheaval of 1848 and the

emergence of Russian sub-state proto-terrorism from 1866 (Verhoeven 2009). The analysis of the graphs' specific details is secondary to that of their relative shapes. From top to bottom in the right column are the fiKorpSV trend graphs for 'terrorist', 'nihilist', 'revolutionary', and 'anarchist', showing co-occurrences among them. The graphs closest to the 'terrorist' profile are 'nihilist' for the 1880s bump (although with a different relative scale) and 'revolutionary' for the 1905 peak of the First Russian Revolution. This makes sense, as the Russian nihilists had been suppressed by the 1890s and were from 1902 replaced by activists of the Party of Socialist-Revolutionaries (SRs). However, the 'terrorist' and 'anarchist' profiles show no strong correlations, indicating that terrorism was at this point not yet understood to also include anarchism.
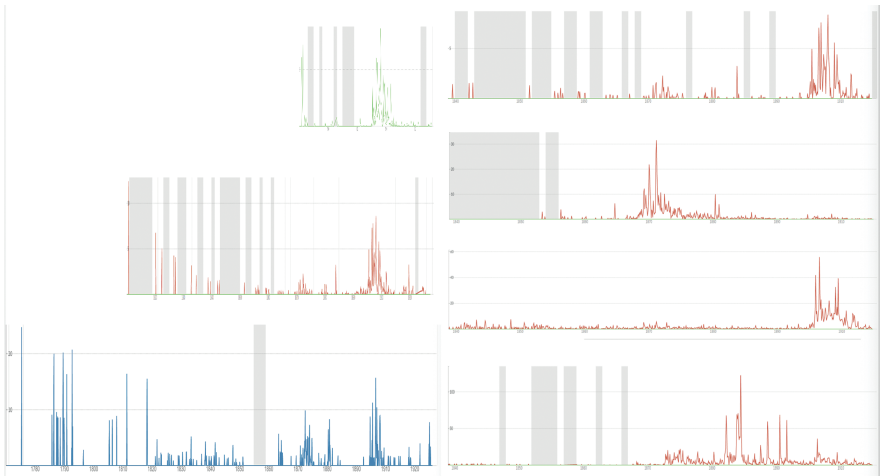
Figure 3: Political militants trend graphs. Left column: *terrorist* for seKorp 1780–1926 (bottom), fiKorpSV 1805–1925 (middle), and *terroristi* for fiKorpFI 1880–1925 (top). Right column (from top): 'terrorist', 'nihilist', 'revolutionary', and 'anarchist' from fiKorpSV for 1848–1920.

A closer reading of the modifiers referring to the classical examples of rebel terrorism in the material reveals that the 'revolutionary', 'nihilistic', and 'socialistic' terrorisms exclusively referred to non-state terrorism, with the exception of 'revolutionary terrorism', which at times also connoted the French Revolution's Reign of Terror and 19th-century French fears of the return of terroristic regimes.

Probably the earliest example before 1900 of non-Russian rebel 'terrorists' are the Fenians. Irish Fenian terrorism appears to a very limited extent (1882–1889) in the material, in reference to local Irish agrarian terrorism in the form of boycotts and murders of English settler farmers ('agrarian murder') and also an urban ter-

rorist campaign with assassinations and bombings (with the agrarian murders as a spin-off). However, these mentions are rare (further discussed in Section 4.4).

Nevertheless, at the beginning of the 1900s other forms of anti-imperial terrorism decidedly entered the terrorism discourse, as evidenced through a spectrum of new nationalities modifiers. One of its earliest manifestations is a Macedonian "band of terrorists, known as the band of dynamitards" that fought Turkish authorities in 1903, using explosives to "gain Europe's attention" (*Åbo Tidning* 1903-10-01). From 1905 'terrorism' became used in relation to both non-socialist and socialist Finnish terrorists, with the 'non-socialist' terrorism referred to as "the budding terrorism in Finland" (1905-08-14) (which receives an in-depth analysis in Section 5). This was followed in 1907 by reports on arrested "Armenian terrorists" in Odessa.

It is noteworthy that among the manifestations of anti-imperial terrorism in the fiKorp material, several are tied to the Russian empire, such as 'Baltic' (1906), 'Latvian' (1907), and 'Grusinian' (1912), although other empires and imperial regions such as Persia, Poland, and Turkey also figure. Also at this time, anti-imperial terrorism appears in East Asia in our dataset. First in 1909 in seKorp and fiKorp descriptions of "Indian terrorists" who renewed their secretive activities and later in 1916 in a unique seKorp mentioning of a female "Chinese terrorist", who took part in the 1911 Xihai Revolution (*Kalmar* 1916-07-15). This woman – a non-Russian, non-socialist revolutionary – indicates how the rise of anti-imperialism contributed to the widening of the notion of sub-state terrorism beyond the Russian context.

4.4 The absent terrorists

For rebel terrorism, an intriguing finding is the contexts in which the word 'terrorism' was *not* used. Several 19th century events frequently described as "terrorism" in historical research do in fact not show up in our word pictures. For example, spectacular terrorist deeds by anarchists in Europe and the USA, as well as anti-imperial separatists in Europe and Asia during the 19th century, were not found to be directly associated with terrorism, neither in Swedish or the Finnish newspapers – attacks which are nowadays seen as constituting parts of a major phase in the history of terrorism, as when David Rapoport writes about "systematic Anarchist efforts to put atrocities in the service of revolution" during "the anarchist wave of rebel terror" (Rapoport 2003: 38). While nihilist terrorists figure prominently in our dataset, there are very few examples of attributions of 'terrorism' to anarchistic activities or, as already discussed above, Fenian activities, which are held to be two of the other major forms of terrorism during the 19th century.

In the case of anarchism, there is only one instance of anarchist terrorism attribution in our dataset in the form of an 1884 article published in two Finnish-Swedish newspapers, reporting that Vienna had been put under a state of siege to make the population safe against "the anarchists' terrorism". Besides that, no other instances of anarchist terrorism appear in the material (although an article in 1880 warns about "anarchic" terrorism in Ireland).

At the same time, we can observe how the terms 'terrorism' and 'terrorist' gradually acquire new meanings and there are indications from 1906 that 'terrorism' became more firmly associated with 'anarchism'. This comes out explicitly, for example, when a Finnish newspaper explains that 'terrorism' and 'anarchism' "are two different concepts", although they "are difficult for a layman to distinguish from each other" (*Nuori Suomi* 1906-02-02). Thus, from then on a more ideologically inclusive 'terrorism' concept is emerging. In 1909, we find that 'terrorism's' history is, so to speak, retrospectively revised accordingly, when an article states that Russian terrorists in the 1890s "had committed anarchist propaganda acts" (*Suomalainen Kansa* 1909-07-13) and in 1910 it was said that 'anarchism' in "everyday speech" had acquired the meaning of perpetrators "of atrocities [*hirmutöiden*]" (*Ilkka* 1910-05-07). By 1912 the conversion appears to have become established, when the famous anarchist tactic of "propaganda by deed" is merged with terrorism, as when militant anarchists were described by *Helsingin Sanomat* as "those 'propaganda by deed' terrorists" (1912-12-10).

5 Trolling for new terrorisms in Finland

To examine how the meanings of 'terrorism' started to broaden outside of the immediate Russian context, we in the following limit our attention to the emergence of the discourse on Finnish terrorism. Here, we leave the indiscriminate trawling "readsearch" method in favour of precision "trolling" (cf. "angling" readsearch in Fridlund 2020) to carefully catch specific emerging meanings in the Finnish context. Thus, we use targeted searches to find Finnish 'terrorism' and 'terrorists' without Finnish pre- or post-modifiers, and also include journals and clandestine newspapers in the fiKorp searches (which were excluded earlier for more equal cross-lingual analysis).

In our searches, references to 'Finnish terrorism' occur in two periods: 1905–1911 and in 1918, the latter referring to Finnish socialists who started the Finnish Civil War in the newly independent state. Our analysis will focus on the discourse during the first period, in order to explore the context of how 'terrorism' came to denote a more ideologically inclusive rebel terrorism closer to the contemporary kind.

5.1 Emergence of restorative and separatist terrorism

From 1899 to 1905, the Russian Grand Duchy of Finland suffered a Russification campaign with increased repression and decrease of its political autonomy. Subsequently, an increasingly violent resistance campaign developed and when the general governor Nikolay Bobrikov was killed by the Finnish nobleman Eugen Schauman in 1904, the duchy had its first act of rebel terrorism. In the assassination's immediate aftermath, no newspapers in our dataset characterized it as 'terrorism'. However, soon afterwards it was interpreted as a sign that Russian rebel terrorism had been appropriated by Finnish nationalists. Yet, the motivations of the Finns were different than the Russian precursors. Schauman's terrorism was not socialist or revolutionary, but "restorative". While directed against the oppressive Russian regime, its foremost aim was not separatism but to restore Finland's earlier autonomy within the empire (Fridlund and Sallamaa 2016: 41; Jensen 2018).

In November 1904, the (non-socialist) Finnish Active Resistance Party (FAM) was founded; modelled on the Russian SR Party, it included a terrorist Combat Organization. Although the Russian tactic and novel institutional form of rebel terrorism thus had arrived, the use of the term 'terrorism' emerged only after a second assassination in February 1905, when FAM sympathizer Lennart Hohenthal shot the Finnish Chancellor of Justice Eliel Soisalon-Soininen.

The official police report in May 1905 did not explicitly call the killing 'terrorism', but it is clear it was viewed as such. According to our results, Finnish newspapers now begun to describe the Soisalon-Soininen killing and similar Finnish acts of political violence in terms of terrorism. According to *Helsingin Sanomat*, the official police report stated that those actively protesting against Russification had incited others to "terrorism and violent acts" (1905-05-05). Although 'terrorism' had rarely been used earlier in relation to similar acts of Finnish political violence, these newspaper accounts basically created a historical narrative about Finnish terrorism, retrospectively.

In *Helsingin Sanomat*, the Soisalon-Soininen assassination was described as born out of previous years-long harassments of Finnish politicians who complied with the Russian regime, especially by Finns close to the Finnish Swedish-language underground journal *Fria Ord*. The journal had revealed "its terrorist purposes by its defamatory accusations and writing in a threatening way", inciting "violence and hatred against officials and supporting revolutionary social democracy, anarchism and terrorism", and by reprinting "defences of murder by Russian terrorists etc., the resistance men have tried to prepare the ground in Finland for such actions" (1905-05-05). The use of the word 'terrorism' was not entirely new in this context, as Finnish supporters of concessions to the Russian oppression had on some earlier occasions used it in reference to threats expressed by their opponents.

The Finnish-nationalist politician Yrjö Sakari Yrjö-Koskinen, a common target of the resistance's harassment and hatred, had already used it in such a way in 1900, in his *Open Letter to my Friends*, where he, as quoted in the police report, accused the resistance of fomenting terrorism, as according to him there had been public and veiled attempts "to implement general terrorism which in my view cannot produce anything but destruction". Even though it is not clear what Yrjö-Koskinen exactly meant by terrorism, the police report on the Soisalon-Soininen killing commended him for daring to "call its actions by their right name: *terrorismi ja hirmu-walta* ['terrorism and reign of terror']" (*Helsingin Sanomat*, 1905-05-05).

Thus, the emergence of Finnish terrorism became framed as a reaction to Russian regime terrorism. The *Karjala* newspaper somewhat later stated that "[w]ith the Russian system, also the concomitant terrorism has been brought to our state's government". Furthermore, Finland was deemed "vulnerable to horrors of terrorism – terrorism that grows and grows", while the Russian regime had tried but failed "to set official terrorism against terrorism" (*Karjala*, 1905-07-23). A similar argument was put forward two months later in an editorial published in several newspapers during the Hohenthal trial. It warned unification-minded "Russian statesmen" that if they knew how much the Gendarmerie military security force in Finland had "deepened the chasm between the Russian government and Finns, then they would immediately demand the abolishment of this institution which only creates a breeding ground for 'separatist terrorism'" (see Figure 2). This referred to the actions of Hohenthal, who had been an informer for the Gendarmerie (*Vaasa* 1905-09-12).

Consequently, these findings show that the meaning of 'terrorism' had by this time become associated with both specifically Finnish strivings as well as a new general motivation. Through this new explicit connection between separatism and terrorism crafted by commentators in the Finnish press, the understanding of rebel terrorism, from then on, was widened from the Russian socialist or revolutionary terrorism to also include anti-imperial separatism. Other parts of the world soon followed, such as when an Indian nationalist in 1907 pointed to "the Russian method" as the most likely one to drive the British out of India (Ker 1917:107).

5.2 Rebel terrorism in the shadow of Russian repression

This new conceptualization of (Finnish) terrorism can also be seen in Sweden, where one of its earliest appearances is almost literally in the actual geographic Gulf of Bothnia. In 1905, the *Kalmar* newspaper reported that a skipper in Northern Finland had found an abandoned shipment of weapons on a rocky islet close to the Swedish border, which might be connected to the "Finnish terrorism, a

child of Bobrikoff and [Russian Minister of the Interior] Plehve, [which] has long striven for association with its Russian kind" (1905-09-18). The statement was later followed by a mentioning in a later article of "the more and more growing crowd of Finnish terrorists" (*Kalmar* 1905-10-25).

The Russification campaign and its 'Years of Oppression' (*sortovuodet*) ended in November 1905 with the Finnish offshoot of the First Russian Revolution, sometimes described as the Finnish Revolution. From now on, the 'terrorism' designation was repeatedly used in the Finnish context. The Constitutionalists party was occasionally accused of 'terrorism' tendencies in their ranks (meaning FAM sympathizers) (see for example, *Uusimaa* 1905-11-10 and *Vaasa* 1905-11-25). In 1906, we can even see that Finnish terrorists were mentioned in a positive, or at least not pejorative way, when a newspaper stated that "[o]ur terrorists during *sortovuodet*, as misguided as their actions could sometimes be, were using violent means to fight for a legal societal system and against its oppressors and destroyers" (*Karjala* 1906-10-28). In 1907 FAM's programme for the first time in public defined Finnish independence and thus separatism as an objective.

However, Russian oppression returned in 1908 and lasted until Finland's independence in 1917. Although the terrorist campaign did not recommence, in 1911 a journal warned about terrorist attacks from below: "Before Bobrikov, no-one in Finland accepted terrorism as a method of liberation struggle and criticized also the Russians for using it. When the oppression by Russians continued, Finns also started to realize that violence from above also produces violence from below and that an unavoidable companion of an oppressive government is terror" (*Keski-Suomen Sanomat*, 1911-09-15). This could, then and now, be read as showing an acceptance of separatism as an ideal and of terrorism as a legitimate tactic against Russian oppression.

In other words, from our results it seems that the cultural-political closeness to the Russian regime and the Russian revolutionary context heavily factored into the development of the discourse on Finnish rebel terrorism – both when it comes to how Finnish nationalists' and separatists' activities were cultivated against the background of the political violence of the Russian revolutionaries and the representatives of the Russian regime, and also how the new Finnish deeds of political violence were framed and evaluated.

6 Conclusions

This study has demonstrated the considerable opportunities for historical analysis afforded by distant reading of national online newspaper corpora through the

Korp interface. A crucial part of the investigation has been the integrative interdisciplinary collaboration between Swedish and Finnish researchers in the history of terrorism and natural language processing, which enabled a complex comparative and contrastive analysis of the historical discourse on terrorism based on both distant and close reading. As we have seen in this chapter, the LT-based automatic linguistic annotations offered by Korp – notably lemmatization and dependency parsing, which enable its "word picture" functionality – together with its sophisticated search abilities add considerable value to this kind of investigation, enabling broad trawling as well as targeted trolling.

As expected, our findings strengthen the earlier hypothesis within history of terrorism that the modern meaning of sub-state terrorism was not widely established in the 19th century. The study also further contributes to the understanding of the historical emergence of terrorism in Europe in at least three ways. Firstly, our results support the supposition that terrorism remained associated with state terror and the Russian context for a long time, but also indicate a great diversity in state character attributions for the later period of the 19th century, as manifested by its presence in a number of national contexts. Secondly, another important finding is the rare occurrence during the 19th century of attributions of terrorism to anarchist and Fenian militants that otherwise figure prominently in the contemporary academic discourse on 19th century terrorism. Although we are not the first to note this, we present new quantitative findings that support this supposition and indicate how such groupings were only in the early 1900s incorporated into the concept of terrorism specifically. Thirdly, we provide a singular exposition of the broadening of terrorism to previously analytically neglected national contexts of anti-colonial separatist terrorism. In this, by turning from trawling to trolling in the form of specifically targeted search methodologies, our investigation yielded detailed novel insights into the domestication of rebel terrorism in Finland. These results indicate that closeness to both the Russian regime and Russian rebels factored into how terrorism in Finland became used specifically in reference to perceived nationalist and separatist activities.

It seems safe to assume that the use of LT resources could contribute further to research on the history of terrorism. Concerning future research, there are now LT methods – for example, topic modelling and (neural) word embedding models – which allow for studies of a fuller range of linguistic expressions of given concepts in vast volumes of text, even in the absence of resources such as Princeton WordNet, which in any case are available only for a few languages.[5] There is a

---

5 In order to be useful for purposes such as the one described here, the vocabulary coverage of such a lexical resource should arguably correspond to a full-sized reference dictionary of a

large wordnet for Finnish (Lindén and Niemi 2014), but not for Swedish, although there are other similar lexical resources for Swedish (see Dannélls, Borin, and Friberg Heppin 2021). When dealing with texts in their entirety rather than words in isolation, in order to see the full picture we should also take into account such linguistic devices as *coreference* – that is when *this phenomenon*, *version*, and *it* all can refer to *terrorism*.

Importantly in our context, the LT methods also apply in multilingual settings, which would allow us to look for word usage correspondences across languages (Ruder, Vulić, and Søgaard 2019). Through wider trawling, one may build a more comprehensive and complex picture of the meanings of terrorism during the 19th and 20th century, and one could go deeper through trolling of terrorism-related nouns such as *attentat* and *dynamitard*, or through the use of diachronic word embeddings to trace conceptual changes in terms over time.

A final given extension of our investigation would be to investigate later periods as well as to seek out the emergence of terrorism in other "discursive transnational bodies of water", including the wider Baltic, Atlantic, and Pacific contexts.

Bibliography

Adesam, Yvonne, Dana Dannélls & Nina Tahmasebi. 2019. Exploring the quality of the digital historical newspaper archive KubHist. *Digital Humanities in the Nordic Countries Conference (DHN) 4*, 9–17.

Backer, Hermanni & Manuel Frias (eds.). 2013. *Planning the Bothnian Sea – key findings of the Plan Bothnia project. Digital edition*. Turku: Finepress Turku.

Borin, Lars, Markus Forsberg & Johan Roxendal. 2012. Korp – the corpus infrastructure of Språkbanken. *International Conference on Language Resources and Evaluation (LREC) 8*, 474–478.

Borin, Lars & Richard Johansson. 2014. Kulturomik: Att spana efter språkliga och kulturella förändringar i digitala textarkiv. Blog Historia i en digital värld.

language, say a minimum of 50,000 entries (Princeton WordNet has more than 150,000 entries). Against this background, lists such as that found at http://globalwordnet.org/resources/wordnets-in-the-world/ are somewhat deceptive. First, many of the 78 links to wordnets provided there are dead. Further, no (easily retrievable) statistics are given even for those wordnets which can be accessed. A fair picture of the state of the art can be gleaned from the Open Multilingual Wordnet page http://compling.hss.ntu.edu.sg/omw/, where data on wordnets for 150 languages are provided, that is about 2% of the world's languages. From this information we can calculate that the average number of words (lemmas) in these 150 wordnets is 10,780, but in true Zipfian fashion the median is much lower: 1,429. A size of 50,000 words is attained only at the 95th percentile, that is, about 8 of these wordnets fulfil the requirement.

Dannélls, Dana, Lars Borin & Karin Friberg Heppin (eds.). 2021. *Swedish FrameNet++: Harmonization, integration, method development and natural language processing applications*. Amsterdam: John Benjamins.

Ditrych, Ondřej. 2011. A genealogy of terrorism in states' discourse. Ph.D. diss., Prague: Charles University.

Ditrych, Ondřej. 2014. *Tracing the discourses of terrorism: Identity, genealogy and the state*. Basingstoke: Palgrave Macmillan.

Erlenbusch, Verena. 2015. Terrorism and revolutionary violence: The emergence of terrorism in the French revolution. *Critical Studies on Terrorism* 8 (2). 193–210.

Fellbaum, Christiane (ed.). 1998. *WordNet: An electronic lexical database*. Cambridge: MIT Press.

Fridlund, Mats. 2018. Digital history 1.5: Historical research between domesticated and paradigmatic digital methods. HumLab Seminar Series, Umeå University, 18 May 2018. https://web.archive.org/web/20190331123529/http://stream.humlab.umu.se/?streamName=digital_history_1_5.

Fridlund, Mats. 2020. Digital history 1.5: A middle way between normal and paradigmatic digital historical research. In Mats Fridlund, Petri Paju and Mila Oiva (eds.), *Digital histories: emergent approaches within the new digital history*, 69–87. Helsinki: Helsinki University Press.

Fridlund, Mats, Daniel Brodén, Leif-Jöran Olsson & Lars Borin. 2020. Trawling the Gulf of Bothnia of news: A big data analysis of the emergence of terrorism in Swedish and Finnish newspapers, 1780–1926. In Costanza Navarretta and Maria Eskevich (eds.), *Selected Papers from the CLARIN Annual Conference 2020, Virtual Edition*, 61–65. (Linköping Electronic Conference Proceedings 180).

Fridlund, Mats, Leif-Jöran Olsson, Daniel Brodén & Lars Borin. 2019. Trawling for terrorists: A big data analysis of conceptual meanings and contexts in Swedish newspapers, 1780–1926. In Melvin Wevers, Mohammed Hasanuzzaman, Gaël Dias, Marten Düring and Adam Jatowt (eds.), *The 5th International Workshop on Computational History (HistoInformatics 2019)*, 30–39.

Fridlund, Mats & Daniel Sallamaa. 2016. Radikale Mittel, gemäßigte Ziele: Repression und Widerstand im Großfürstentum Finnland. *Osteuropa* 66 (4). 35–47.

Graham, Shawn, Ian Milligan & Scott Weingart. 2016. *Exploring big historical data: The historian's macroscope*. London: Imperial College Press.

Ifversen, Jan. 2011. About key concepts and how to study them. *Contributions to the History of Concepts* 6 (1). 65–88.

Jackson, Richard. 2008. The ghosts of state terror: Knowledge, politics and terrorism studies. *Critical Studies on Terrorism* 1 (3). 377–392.

Jensen, Richard Bach. 2018. The 1904 assassination of governor general Bobrikov: Tyrannicide, anarchism, and the expanding scope of "terrorism". *Terrorism and Political Violence* 30 (5). 828–843.

Karsvall, Olof & Lars Borin. 2018. SDHK meets NER: Linking place names with medieval charters and historical maps. *Digital Humanities in the Nordic Countries Conference (DHN) 3*, 38–50.

Ker, James Campbell. 1917. *Political trouble in India*. Calcutta: Superintendent Government Printing, India.

Kujala, Antti. 1992. Finnish radicals and the Russian revolutionary movement, 1899–1907. *Revolutionary Russia* 5 (2). 172–192.

Law, Randall D. 2009. *Terrorism: A history*. Cambridge: Polity Press.

Lindén, Krister & Jyrki Niemi. 2014. Is it possible to create a very large wordnet in 100 days? An evaluation. *Language Resources and Evaluation* 48 (2). 191–201.

Miller, Martin A. 2013. *The foundations of modern terrorism: State, society and the dynamics of political violence*. Cambridge: Cambridge University Press.

NLF. 2011a. The Finnish Sub-corpus of the Newspaper and Periodical Corpus of the National Library of Finland, Kielipankki Version.

NLF. 2011b. The Swedish Sub-corpus of the Newspaper and Periodical Corpus of the National Library of Finland, Kielipankki Version.

Paju, Petri, Mila Oiva & Mats Fridlund. 2020. Digital and distant histories: Emergent approaches within the new digital history. In Mats Fridlund, Petri Paju and Mila Oiva (eds.), *Digital histories: Emergent approaches within the new digital history*, 3–18. Helsinki: Helsinki University Press.

Pettersson, Eva & Lars Borin. 2022. Swedish Diachronic Corpus. In Darja Fišer and Andreas Witt (eds.), *CLARIN. The infrastructure for language resources.* Berlin: De Gruyter.

Rapoport, David C.. 2003. The four waves of rebel terror and September 11. In Charles W. Kegley Jr. (ed.), *The new global terrorism: Characteristics, causes, controls*, 36–52. Upper Saddle River: Prentice Hall.

Ruder, Sebastian, Ivan Vulić & Anders Søgaard. 2019. A survey of cross-lingual word embedding models. *Journal of Artificial Intelligence Research* 65: 569–631.

Sageman, Marc. 2017. *Turning to political violence: The emergence of terrorism*. Philadelphia: University of Pennsylvania Press.

Salmi, Hannu, Petri Paju, Heli Rantala, Asko Nivala, Aleksi Vesanto & Filip Ginter. 2013. The reuse of texts in Finnish newspapers and journals, 1771–1920: A digital humanities perspective. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 54 (1). 14–28.

Tahmasebi, Nina, Lars Borin, Gabriele Capannini, Devdatt Dubhashi, Peter Exner, Markus Forsberg, Gerhard Gossen, Fredrik Johansson, Richard Johansson, Mikael Kågebäck, Olof Mogren, Pierre Nugues & Thomas Risse. 2015. Visions and open challenges for a knowledge-based culturomics. *International Journal on Digital Libraries* 15 (2–4). 169–187.

Tahmasebi, Nina, Niklas Hagen, Daniel Brodén & Mats Malm. 2019. A convergence of methodologies: Notes on data-intensive humanities research. *Digital Humanities in the Nordic Countries Conference (DHN)* 4, 437–449.

Tangherlini, Timothy R. & Peter Leonard. 2013. Trawling in the Sea of the Great Unread: Sub-corpus topic modeling and Humanities research. *Poetics* 41 (6). 725–749.

Verhoeven, Claudia. 2004. April 4, 1866: The Karakozov Case and the Making of Revolutionary Terrorism. Ph.D. diss., University of California, Los Angeles.

Verhoeven, Claudia. 2009. *The odd man Karakozov: Imperial Russia, modernity, and the birth of terrorism*. Ithaca: Cornell University Press.

Viklund, Jon & Lars Borin. 2016. How can big data help us study rhetorical history? In Koenraad De Smedt (ed.), *Selected Papers from the CLARIN Annual Conference, October 14–16, 2015, Wrocław, Poland*, 79–93. (Linköping Electronic Conference Proceedings 123).

Weller, Toni (ed.). 2013. *History in the digital age*. London: Routledge.