

Trawling for Terrorists: A Big Data Analysis of Conceptual Meanings and Contexts in Swedish Newspapers, 1780–1926

Mats Fridlund¹²⁵, Leif-Jöran Olsson³⁴, Daniel Brodén¹²⁴, and Lars Borin³⁴

¹ Centre for Digital Humanities, University of Gothenburg
cdh@lir.gu.se

² Dept. of History of Ideas, Literature and Religion, University of Gothenburg
mats.fridlund|daniel.broden@lir.gu.se

³ Språkbanken Text, Dept. of Swedish, University of Gothenburg
leif-joran.olsson|lars.borin|sb-info@svenska.gu.se

⁴ Swe-Clarín

info@sweclarin.se

⁵ Max Planck Institute for the History of Science, Berlin
mfridlund@mpiwg-berlin.mpg.de

Abstract

The conceptual history of terrorism has to a significant extent been studied through canonical texts or historical key figures or organisations. However, through the increasing digitization of text materials conventional research questions can now be approached from new angles or established results verified on the basis of exhaustive collections of data, rather than limited samples. Specifically, we are interested in evaluating and expanding on prior research claims regarding the meanings and contexts associated with the concepts terrorism and terrorist up until the twentieth century in a Swedish context. The investigation is guided by the following research questions: What historical meanings of the concept of terrorism were expressed in the Swedish newspaper discourse? What social and ideological contexts and violent political practices was the concept primarily associated with before the First World War?

1 Introduction

1.1 Background

History like the other disciplines of the humanities are entering the age of big data. Digital technologies and mass digitization of historical documents provide novel ways for researchers to analyze textual materials. New research questions can be answered, old ones addressed from new angles and established results verified based on big data collections, rather than limited samples or case studies. Now, we can process huge data collections using language technology (LT), allowing us to search for abstract patterns in massive amounts of text and track conceptual changes over time in broad outline as well as in minute detail. However, there are significant methodological challenges to be overcome. High-profile big data initiatives have been criticized for, among other things, a lack of linguistic sophistication of the basic text processing (see [Zimmer, 2013](#) and [Tahmasebi et al., 2015](#)) and a lack of awareness of questions of validity and representativity of the data sets, that is whether they actually support the claims being made (see [Pechenick et al., 2015](#)). [Tahmasebi et al. \(2019b, 439\)](#) state that:

typical digital humanities projects are conducted with either a strong data science or humanities bias. The data science projects, on the one hand, are often conducted with a computer science, math, or language technology perspective where the interpretation and the understanding of the research questions are sacrificed at the expense of mining techniques and large quantities of data. The humanities projects, on the other hand, are often conducted on smaller scale data using methods that may not be the best suited for the problem, or data, at hand

Historical research on the emergence since the late 18th century of terrorism as a systematic political practice primarily consists of qualitative studies, focusing on analysing primary and secondary source texts describing the activities of central historical figures and organisations using terrorist practices (see for example Land, 2008; Verhoeven, 2009; Sageman, 2017). Only rarely does one find that the primary research on the historical discourse or on descriptions of terrorism is supplemented by quantitative studies using text mining of online newspaper archives for occurrences of significant conceptual terms such as *terrorism* or *terrorists* (see Ditrych, 2011, 2014). So far there have been no historical studies systematically using quantitative methods to study the emergence of terrorism as a concept and as a phenomenon in its historical contexts.

1.2 History of Terrorism meets Language Technology

The aim of this paper is to explore the usefulness of large-scale corpus studies in the history of terrorism by showing how the historical concept of terrorism can be investigated using LT-based big-data methods. In doing this the study distinguishes itself by combining LT expertise with humanities domain expertise, notably in history of terrorism (Fridlund, 2011, 2012a,b, 2015). Specifically, we are interested in evaluating and expanding on prior research claims regarding the meanings and contexts associated with the concept up until the 20th century in a Swedish context.

The study is carried out with support of the e-research infrastructure CLARIN (Common Language Resources and Technology Infrastructure), a European Research Infrastructure Consortium (ERIC) which aims at making various language-based materials available as primary research data to the humanities and social sciences (HSS) with the aid of state-of-the-art language and speech processing tools and language resources. The research is part of an initiative launched by the Swedish CLARIN node, Swe-Clarín, consisting in pilot projects where HSS scholars are working together with researchers in natural language processing and an e-science infrastructure unit (Swe-Clarín/Språkbanken Text) on designing, developing and evaluating LT-based e-science tools for HSS. Other such pilot projects have addressed conceptual changes in the vocabulary of rhetorics (Viklund and Borin, 2016) and automatic named-entity recognition in Swedish medieval charters (Karsvall and Borin, 2018).

The analysis uses the Swedish newspaper corpus Kubhist and the corpus search tool Korp. Both Kubhist and Korp are 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-Clarín, the national Swedish CLARIN ERIC organization.

Korp (Borin et al., 2012) is an advanced corpus search tool with modular design and a flexible web user interface. Although primarily designed to fulfill the research needs of linguists, Korp has turned out to be a useful research tool also for addressing other kinds of questions in the humanities and social sciences, not least because it provides access to about 15 billion words of Swedish text from various historical periods, through a sophisticated online search interface. The interface lets you make both simple word searches and more advanced combined queries, based on automatic linguistic annotations. The results can be presented in different ways. Firstly, as a *contextual hit list* or *KWIC* (keyword in context). Secondly, as statistical data with occurrence frequencies in sub-corpora, which, among other things, let you create a *trend graph* with relative frequencies plotted over time for text words, lemmas

(dictionary headwords), or other linguistic items. Thirdly, as a so-called *word picture*, which presents statistically prominent fillers of selected syntactic dependency relations of a word, e.g. typical subjects and objects of a verb, and nominal premodifiers (typically adjectives) and postmodifiers (prepositional phrases or main verbs of relative clauses).¹ Importantly, the word picture can be used as an abstract topical map that guides you to a closer reading of the corpus. Korp further supports effortless navigation back and forth between the distant-reading views afforded by the statistics, trend-graph and word-picture views on the one hand, and the KWIC view allowing close reading of individual hits in their context.

Kubhist is a large corpus of digitized historical newspapers from late 18th to early 20th century provided by the Swedish National Library. The Kubhist corpus currently contains about 1 billion words, and the next version, available and planned for inclusion in Korp in the near future, comprises on the order of 5.5 billion words. For a more detailed description of Kubhist, see [Tahmasebi et al. \(2019a\)](#). Although Kubhist is relatively small compared to the Google Books dataset, the corpus distinguishes itself from Google Books and other digital historical newspaper projects such as *impresso*, *Europeana*, and *NewsEye* by being linguistically annotated on several levels (lexical, morphological, lexical-semantic, syntactic, named entities, etc.). In particular, the annotation tools draw on high-quality lexical resources (historical as well as modern). This arguably compensates for the smaller size ([Borin and Johansson, 2014](#); [Tahmasebi et al., 2015, 2019a](#)). Moreover, Kubhist is far larger than any material previously used for investigating the Swedish press discourse during the 18th and 19th centuries.

Our analysis focuses on the period 1780–1926 that is chosen mainly for two reasons. Firstly, the choice is pragmatic. Although Kubhist covers the years 1749–1926, the newspaper material is not complete until 1780 and the corpus ends in 1926 due to copyright restrictions. Secondly, we took 1780 as the starting year because it is almost ten years before the start of the French Revolution, whose Reign of Terror is a key development in the history of terrorism. This period lets us explore the conceptual development of terrorism since its first use in France in 1794 to describe a new political state practice ([Erlenbusch, 2015](#)) and its subsequent development during the late 19th century when its meaning expanded to include political violence used by substate actors in Russia before it further broadened during the early 20th century closer to the contemporary usage describing various substate political violence with diverse ideological motivations. The meanings of terrorism that we investigate here is what kind of actors and violent political practices that were described as terrorists and terrorism and what particular ideologies that the terrorists allegedly adhered to and how these changed over time, while the contexts of terrorism focus on investigating the various physical and social places and spaces in where terrorism and terrorists occurred.

Preliminary results from historical research ([Fridlund, 2018](#)) on the use of ‘terrorism’ in Swedish newspaper materials 1848–1920 have indicated that the dominant modern meaning of terrorism, to describe a violent practice adopted by a wide range of substate political militants with diverse ideological motivations, was more widely established only during the 20th century. The traditionally accepted view of the emergence of substate terrorism as a systematic political practice is that it was first introduced by Russian social revolutionaries in the late 1870s and then by the end of the 19th century spread internationally by being appropriated by revolutionaries, anarchists and colonial freedom fighters in Europe, North America and Asia. However, this new research is indicating that although terrorist tactics were used by 19th century anarchists and colonial militants, the concept’s meaning was initially more specific and restricted and only used more widely outside of the Russian context in the 20th century. To study the ideologies associated with terrorism the investigation is thus guided by the following research ques-

¹Note that these are pre- and post-modifiers in the *linguistic* sense, drawing on the automatic syntactic (dependency) annotations added to Språkbanken Text’s corpora as part of the corpus import process. This is quite different from the word n-gram window based collocation measures normally encountered in related work, which crucially rely on the investigated language having a mainly fixed word order (and mainly short phrases). Using dependency relations as the basis for calculating collocations largely removes such limitations.

tions: *What historical meanings of the concept of terrorism were expressed in the Swedish newspaper discourse? What social and political contexts was the concept primarily associated with?*

2 Analysis

Terrorism is part of a longer historical tradition of wider practices of political violence exercised by both state and non-state actors. To include this wider context of political violence we in the searches conducted in Korp combined the queries for *terrorist* (259 hits) and *terrorism* (570 hits) with closely associated terms used for actors using political and non-political violence such as *soldat* ‘soldier’ (69,979 hits), *socialist* (10,600 hits), *anarkist* ‘anarchist’ (3,028 hits), *bandit* (3,907 hits), *nihilist* (1,660 hits), *revolutionär* ‘revolutionary’ (noun: 1,285 hits; adjective: 8,024 hits), and *härskare* ‘ruler’ (569 hits). The terms ‘nihilists’ and ‘nihilism’ were often used as synonyms for Russian socialist revolutionaries and their ideologies, as well as for adherents of various philosophical and value-negating thought systems. Among these we looked more in detail at ‘anarchist’, ‘nihilist’, and ‘revolutionary’ together with the corresponding ideological words ‘anarchism’, ‘nihilism’ and ‘revolutionism’. Furthermore, we studied the more general word *attentat*, which although prominently used in French, German and Swedish especially for public assassination attempts does not have an exact English equivalent and is often merely translated as ‘attempt’. It is mostly used to refer to violent political deeds in general and thus yields a much larger number of hits (8,801) compared to those of ‘terrorist’ (259) or ‘anarchist’ (3,028).

2.1 Trend Graphs

Figure 1 shows trend graphs for some of the central words involved in expressing the concept terrorism over the period 1780–1914, and in Figure 2, we zoom in on the end of this period, ca. 1840–1914. The reason for this focus on a particular period is that we know from the historical record that the modern practice of substate terrorism emerges from 1866 onwards with a number of violent political propagandistic deeds by Russian social revolutionaries and self-proclaimed terrorists such as the first terrorist organisation Russian Narodnaya Volya.

The top trend graphs in Figure 1 and Figure 2 display the occurrences for *terrorism* and *terrorist* and show that this distant-reading of terrorism in Sweden before the First World War appears to reflect some – but not all – well known facts of the known historical record of terrorism. Although the details of what the actual occurrences correspond to have to wait for closer readings of results using KWIC, and possibly also of the actual newspaper articles, one can at this stage surmise that the hits in the 1790s correspond to the Reign of Terror during the French revolution, the outbreak of Russian nihilist terrorism in the 1870s and 1880s as well as the new wave of Russian revolutionary terrorism before, during and after the 1905 Russian revolution. An interesting finding is the bunching around the late 1840s and early 1850s that might be connected to the European revolutions of 1848. There are no well known incidences of state or substate terrorist attempts during this period so it will be interesting to see from KWIC whether that could be referring to the state terrorism of the French revolution or more likely fears of new nonstate revolutionary terrorism in the vein of the French revolution. Overall the *terrorism/terrorist* trend graph gives confidence that the close reading should produce further relevant and interesting results.

The next step is to take a look at the trend graphs for 1840–1914 describing the various political actors such as anarchists and nihilists that have traditionally been seen as making up a large part of the terrorists of the period. What the diagram shows are that although there are co-occurrences between terrorists and the other perceived ideological militants there are no strong correlations. Thus we need to go into KWIC and close reading of the articles to see whether those close readings produced stronger correlations. What this result however does indicate is that there are no strong correlations to especially

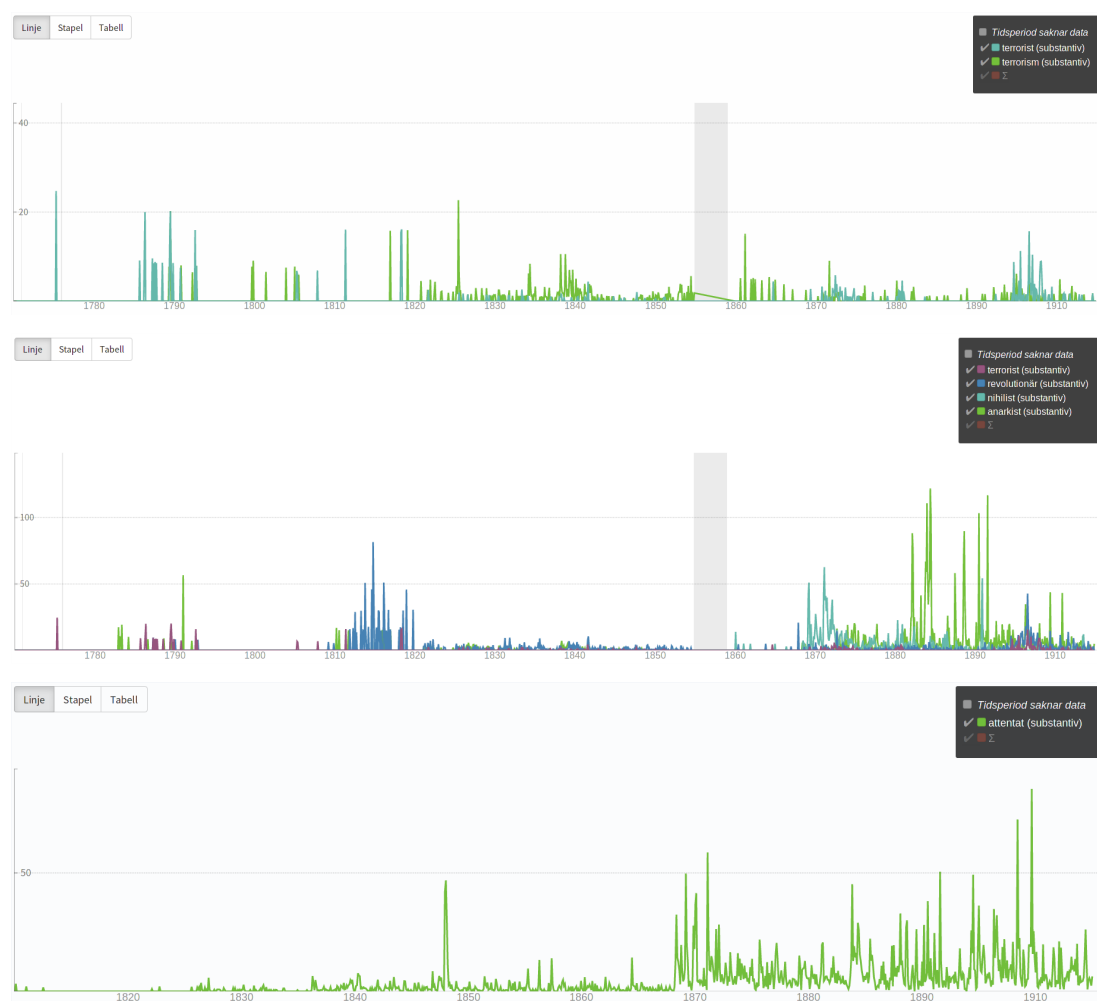


Figure 1: Trend graphs for *terrorist/terrorism* (top), *terrorist/anarkist/nihilist/revolutionär (noun)* (middle), and *attentat* (bottom) for the period 1780–1914

anarchism, which supports and strengthens the hypothesis that terrorism have not reached its modern meaning during this period.

2.2 Word Pictures

Figure 3 shows two examples of word pictures, for the noun lemmas *terrorist* and *terrorism*, with their prominent pre- and post-modifiers.²

Looking at the word pictures and starting with *attentat*, most of the postmodifiers of this word concern who it was directed against (*mot* ‘against’) and although the most prominent one is ‘against life’ (*mot lif*, 154) following that are a number of attributes singling out *attentat* against heads of state and rulers such as against the ‘emperor’, ‘czar’, ‘king’, ‘sultan’, ‘president’, ‘queen’ as well as named

²Other word pictures discussed here are not shown for reasons of lack of space.

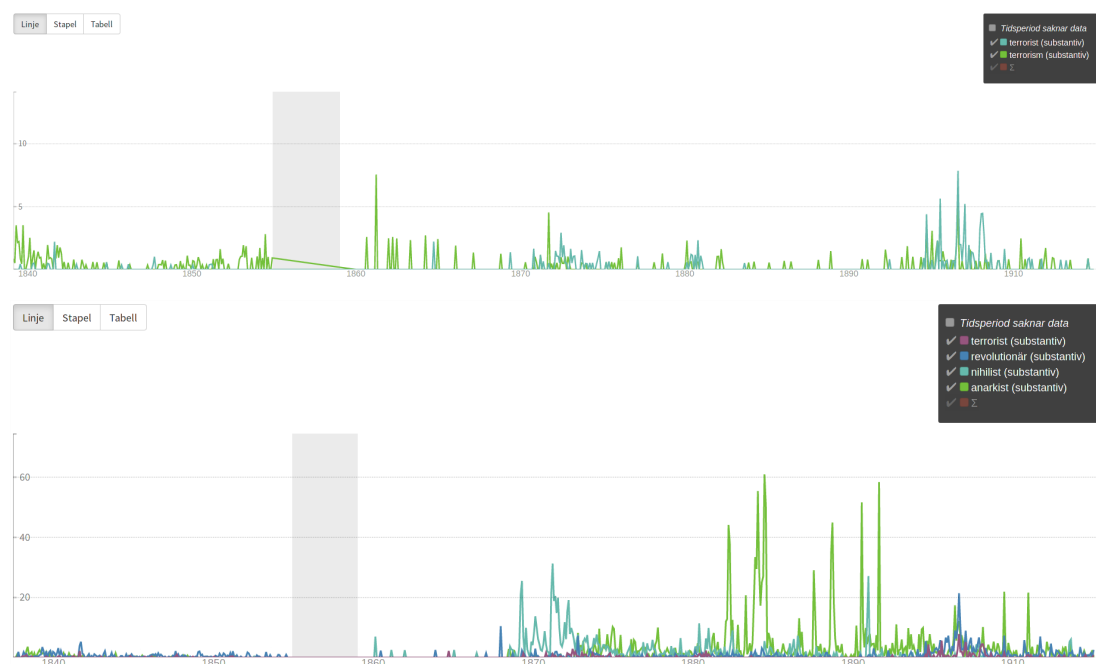


Figure 2: Trend graph for *terrorist/terrorism* (top) and *terrorist/anarkist/nihilist/revolutionär* (*noun*) (bottom) 1849–1914

individual rulers that suffered terrorist as well as non-terrorist attempts on their lives such as *Wilhelm* (39), *Napoleon* (24), *Alexander* (21), *Alfonso* (20), *Stolypin* (16), *Bobrikoff* (12), *Faure* (14), *Victoria* (10), *Estrup* (10), and *Garfield* (10). The only identifiable individual in this set who was not a political ruler is *Hammar* (10). The Swedish industrialist John Hammar in 1909 had one of history's early letter bombs sent to him which blew off two of his fingers. This was a false flag political *attentat* billed as motivated by anarchism but was actually an act of personal revenge against Hammar by a disgruntled engineer.

Besides being directed against people *attentats* were also directed against non-living entities such as infrastructure and ideals, against 'railway trains' (in total 49), 'freedom' (12) and 'freedom of the press' (*tryckfrihet*, 10). Especially the railway *attentats* are most likely describing terrorist deeds as the Russian terrorist group Narodnaya Volya in 1879 directed a prominent spectacular bomb attack against a railway train of the czar.

When we go to the word pictures for *terrorist* and *terrorism* (Figure 3), there are a number of salient findings pointing to several contexts where the *terrorist* figured during the period.

We start by investigating the spatial contexts – both where terrorism occurred as well as where terrorists came from – by looking at the national or ethnic attributes of *terrorism* and *terrorist*. The most common are 'Russian' (*rysk*, 15 occurrences) which points to the fact that the most well known terrorists were active in or came from Russia where they were fighting against state despotism. The other nationalities attributed are 'Finnish' (*finsk*, 3), 'Polish' (*polsta* should be *polska*, 2), 'Hungarian' (*ungersk* 2), 'Irish' (*irländsk*, 1) and 'Chinese' (*kinesisk*, 1). This is an interesting result as it appears to be a mix of state terrorism where 'Polish' and 'Hungarian terrorism' most likely point to state terrorism, in Poland by the Russian government which ruled over the major part of Poland during most of the period under scrutiny here. The non-state terrorism appears to be represented especially by Russia and Ireland,

Preposition	Attribut	terrorist	Efterställt Attribut
1. af	48	1. rysk	15
2. hos	4	2. ifrig	4
3. bland	2	3. utsmigna	2
4. med	4	4. kvionlige	2
5. enligt	1	5. håktade	2
6. öfver	2	6. ffterlomne	2
7. mot	2	7. frigifna	2
8. åt	1	8. verklig	4
9. sedan	1	9. finsk	3
10. på	2	10. revolutionär	2
11. efter	1	11. verkliga	2
12. från	2	12. kvinnlig	2
13. till	2	13. kvinnlig ²	2
14. om	1	14. wore	2
15. av	1	15. själf	2
		16. absolutistisk	1
		17. ka	2
		18. namnkunnig	1
		19. dödsdömd	1
		20. ung	2
		21. utpräglad	1
		22. industriell	1
		23. irländsk	1
		24. bekant	1
		25. sjelfva	2
		26. monarkisk	1
		27. kinesisk	1
		28. farlig	1
		29. framstående	1
		30. riktig	1
		31. militär	1
		1. tillfredstälde	2
		2. från minskov	2
		3. Skte	2
		4. afiossade	2
		5. iätt	2
		6. voro	2
		7. i hufvudstaden	2
		8. hos Antonia	1
		9. samhällsställning	1
		10. till grad	1
		11. i åker	1
		12. vara	2
		13. polisman	1
		14. skola ²	2
		15. uppträda	1
		16. orsak	1
		17. döda	1
		18. för dag	1
		19. tro	1
		20. skola	1
		21. i del	1
		22. i Frankrike	1
		23. stad	1
		24. i Petersburg	1
		25. ha	1

Preposition	Attribut	terrorism	Efterställt Attribut
1. af	58	1. hvita	10
2. mot	12	2. forfäliga	4
3. genom	15	3. revolutionär	5
4. under	14	4. klatorisk	2
5. öfver	10	5. absolutistiska	2
6. nedifrån	3	6. widgiriga	2
7. oaktadt	4	7. katkowfka	2
8. med	10	8. frukreaktiouär	2
9. om	6	9. stywärdaste	2
10. på	8	10. röd	4
11. till	14	11. dim	2
12. för	12	12. waldfammas	2
13. öfwer	2	13. militarisk	2
14. bakom	1	14. inqvisitoriska	2
15. från	4	15. förfärlig	3
16. före	1	16. polsta	2
17. inför	1	17. .:	2
18. av	3	18. gifna	2
19. mellan	1	19. fullkomlig	3
20. till och med	1	20. ig	2
21. till och med	1	21. formlig	2
22. åt	1	22. ohygglig	2
23. efter	1	23. socialistisk	2
24. i	9	24. fullständig	3
		25. autokratisk	1
		26. ungersk	2
		27. svär	2
		28. militär	2
		29. politisk	2
		30. sammansvuren	1
		1. utöfvades	4
		2. till jesuiterne	2
		3. i fini	2
		4. sedan deig	2
		5. i meniugsfrågor	2
		6. kunnen	2
		7. utölväs	2
		8. utöfvar	2
		9. sväfvar	2
		10. emot embetsman	2
		11. utöfva	2
		12. utöfvat	2
		13. hrr	2
		14. i Frankrike	3
		15. hindra	2
		16. vara	3
		17. nedifrån despot	1
		18. i Paris	2
		19. gentemot arbetare	1
		20. förtrycka	1
		21. misstänksamhet	1
		22. mot minoritet	1
		23. mot församling	1
		24. folkvälde	1
		25. mot tänkande	1
		26. till tro	1
		27. inom gebit	1
		28. till Napoleon	1
		29. i Kroatien	1
		30. bota	1

Figure 3: Word pictures showing pre- and post-modifiers for *terrorist* (left) and *terrorism* (right)

as there are many known terrorists of these nationalities, especially the Irish-American Fenians were among the pioneering terrorist groupings of the 1880s. More surprising is the prominence attributed to Finnish and Chinese terrorists. This probably refers to either a Finnish terrorist wave 1904–06 or the ‘red’ and ‘white terror’ of the Finnish Civil War in 1918, something which is supported by ‘white’ and ‘red terrorism’ having 10 respectively 4 occurrences although they could also be referring to French discourse and fears of renewed red revolutionary terrorism as well as occurrences of post-revolutionary white state terrorism in the early 19th century. The Chinese terrorism might be the wave of terrorist ‘assassinationism’ that China suffered in 1904-05.

The most interesting result is however a negative finding, the nationalities *not* attributed to terrorists. It is well known that there were a number of spectacular anarchist terrorist deeds during the period in Germany, Italy, Spain, USA and the UK as well as anti-colonial activists, nationalists and separatists in India and the Ottoman Empire, but none of these nationalities are attributed to the terrorists mentioned in the Swedish newspapers. This is another strong indication that ‘terrorists’ were primarily seen as equivalent with Russian militants. Also the postmodifiers support this as the national locations mentioned for terrorists and terrorism are ‘in France’ (4), ‘in Paris’ (2), ‘in [St] Petersburg’ (1), ‘in Russia’ (1), ‘in Ireland’ (1), ‘in Croatia’ (1) and ‘abroad’ (*i utland*, 1). Here ‘France’ most likely refers to the terrorists of the French Revolution. Furthermore, the occurrence of ‘Paris’ and ‘Petersburg’ emphasizes, and is also shown by postmodifiers such as ‘in the capital’ and ‘city’, that terrorism is perceived as a primarily urban phenomenon, although we know from previous historical research that Irish agrarian political violence and intimidation was also referred to as terrorism.

While the urban categorization is indirect, the results also indicate a direct categorization of terrorists in the form of attributes of ideological or political affiliation such as ‘revolutionary’, ‘monarchical’,

‘autocratic’, ‘oligarchic’, ‘dictatorial’, ‘military’, ‘inquisitorial’, ‘socialistic’, ‘nihilistic’ and ‘industrial’ terrorists. ‘Revolutionary’, ‘nihilistic’, and ‘socialistic’ most likely refer to nonstate terrorism although revolutionary could also be the new revolutionary French state’s terrorism of Robespierre’s Terror after the French revolution and socialistic terrorism that of Lenin after the 1917 Russian revolution. Many of the other characterizations appear to represent state terrorism which makes sense in that a prominent form of state terrorism during the period was the use of terror by despotic monarchical governments and military forces, as foreign occupiers or against their own populations. This makes clear that states were also seen and described as terrorists. The expression ‘industrial terrorist’ might refer to violence and intimidation exercised by laborers and labor unions against employers or against strikebreakers during strikes. Although Sweden’s first terrorist deed in the form of the so called Amalthea bombing in Malmö harbor in 1908 was directed against foreign strikebreakers, and could be seen as ‘industrial terrorism’. However, we know from ongoing qualitative research that this kind of terrorism rarely went as far as lethal violence but was mostly confined to harassment and threats of violence. Also this is surprising that it got such a prominent attribution unless it is connected to the Amalthea bombing mentioned above. Overall what is striking after the reading is that when ‘terrorism’ is used it appears to be pointing primarily to state terrorism while ‘terrorist’ appears to be used to denote nonstate terrorism.

This study’s most noteworthy finding comes out of the several synonyms of ‘female’ (*kvinnlig*, 6) attributed to terrorists. This is especially interesting in comparison with similar word pictures for anarchists, revolutionaries and Fenians where we can find no such similar prominent attribution, if it is mentioned at all. This also supports that ‘terrorists’ were primarily seen as Russians as several of the prominent and leading Russian terrorists were women, such as the first modern terrorists Vera Zasulich and Sophia Petrovskaya, one of the leaders of the world’s first terrorist group. This was also explicitly noted by other public commentators, as can be seen in the Swedish national encyclopedia’s article on political nihilism where it says that “[i] ögonen fallande är mängden af bildade unga ryskor, som med dödsförakt gå nihilismens ärenden och gifva den dess exaltation, dess martyrglöd.” [notable are the many young educated Russian women who without fear of death engage in nihilism and give it its exaltation, its glow of martyrdom] (*Nordisk Familjebok, Militärkonventioner–Nådaval* [1887], s.v. *Nihilism*) In line with and supporting this is that ‘young’ (*ung*, 2) is another prominent attribution of the terrorists.

A last prominent terrorist context is that of the legislative sphere, where we find a relatively large number (10) of verbs such as ‘execute’, ‘hang’, ‘want for questioning’ (*efterlysa*) and ‘arrest’ referring to the punitive legislative measures. This makes good sense as most terrorism then and now involved violence and thus often constituted a crime and the government’s repression of terrorists and the following trials were a prominent feature in this struggle as well as of the militants’ way of communicating their political messages publicly.

3 Conclusions and Future Work

This small study confirms to some extent the common wisdom that the meanings and connotations of the concept of terrorism have changed over time. In this proof-of-concept study we have focused on the distant reading or bird’s-eye view afforded by the Korp interface, but not followed and confirmed these changes in detail, which the interface also allows for through closer reading of its KWIC view, but we can nevertheless say that the material indicates unequivocally that there have been changes over the last centuries. In doing this the study like many similar quantitative text mining studies in the digital humanities confirms what we already knew from the historical record.

Even with this large amount of text, many queries returned only a few hundred hits, which may not allow for very solid generalizations. The next version of Kubhist will be five times as large and also have better OCR quality, a factor that is known to influence search accuracy ([Jarlbrink and Snickars, 2017](#)), and we are looking forward to continuing this work using this larger dataset. With this expanded

dataset it could also be possible to test the usefulness of diachronic word embeddings or other emerging semantic technologies for studying the change in meanings of terms related to terrorism over time, although we do need less data-hungry quantitative methods in any case, if we are to work on more languages and more varied source texts (Tahmasebi, 2018).

However, the distant reading results already go beyond common historical wisdom in pointing to that the diversity of the meaning of terrorism in the period not only includes the classical case of the terrorism of the French revolution but also as especially demonstrated by the state character attributes in the word pictures of *terrorism* and *terrorist* for the later period during the 19th century that appears to have prominent attributions of state terrorism. Additionally, the hypothesis that the modern meaning of substate terrorism was not yet established in this period but primarily restricted to Russian terrorism is also strengthened by the results.

Of course, it would be interesting to also study how the meanings and connotations of the concept of terrorism have changed in Swedish newspapers in the later 20th century perhaps even going into the contemporary period and its increased topicality and global political relevance after the 2001 terrorist attacks in the USA, but this must be left for future work. At present, IPR restrictions make such an investigation impossible, but the National Library of Sweden together with Språkbanken Text and other stakeholder institutions are conducting negotiations with collection societies and other representatives of copyright holders in order to facilitate research involving text processing such as that used for the present investigation. Then Swedish digital historians making arguments using big data analysis could contribute not only to the history of the emergence of terrorism but also to the investigation of its stability and salient presence in the history of the very near present.

Acknowledgements

The work presented here has been partly supported by an infrastructure grant to Språkbanken Text and the Center for Digital Humanities, University of Gothenburg, for contributing to building and operating a national e-infrastructure funded jointly by the Swedish Research Council (under contract no. 2017-00626) and the participating institutions.

References

- Borin, L., Forsberg, M., and Roxendal, J. (2012). Korp – the corpus infrastructure of Språkbanken. In *Proceedings of LREC 2012*, pages 474–478, Istanbul. ELRA.
- Borin, L. and Johansson, R. (2014). Kulturomik: Att spana efter språkliga och kulturella förändringar i digitala textarkiv. In Parland-von Essen, J. and Nyberg, K., editors, *Historia i en digital värld*.
- Ditrych, O. (2011). *A Genealogy of Terrorism in States' Discourse*. Charles University, Prague.
- Ditrych, O. (2014). *Tracing the Discourses of Terrorism: Identity, Genealogy and the State*. Palgrave Macmillan, Basingstoke.
- Erlenbusch, V. (2015). Terrorism and revolutionary violence: The emergence of terrorism in the French revolution. *Critical Studies on Terrorism*, 8(2):193–210.
- Fridlund, M. (2011). Buckets, bollards and bombs: Towards subject histories of technologies and terrors. *History and Technology*, 27(4):391–416.

- Fridlund, M. (2012a). Affording terrorism: Idealists and materialities in the emergence of modern terrorism. In Taylor, M. and Currie, P., editors, *Terrorism and Affordance*, pages 73–92. Continuum, London & New York.
- Fridlund, M. (2012b). Det nya gränsöverskridande våldet: Terroristen och den tidiga globaliseringen. In Svensson, B. and Wallethe, A., editors, *Individer i rörelse: Kulturhistoria i 1880-talets Sverige*, pages 194–230. Makadam, Göteborg.
- Fridlund, M. (2015). Motståndets materialitet: oppositionella ting, tekniker och kroppar under Finlands ofärdsår. In Wormbs, N. and Kaiserfeld, T., editors, *Med varm hand: Texter tillägnade Arne Kaijser*, pages 53–84. KTH Royal Institute of Technology, Stockholm.
- Fridlund, M. (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.
- Jarlbrink, J. and Snickars, P. (2017). Cultural heritage as digital noise: Nineteenth century newspapers in the digital archive. *Journal of Documentation*, 73(6):1228–1243.
- Karsvall, O. and Borin, L. (2018). SDHK meets NER: Linking place names with medieval charters and historical maps. In *Proceedings of DHN2018*, pages 38–50, Aachen. CEUR-ws.org.
- Land, I., editor (2008). *Enemies of Humanity: The Nineteenth-Century War on Terrorism*. Palgrave Macmillan, New York.
- Pechenick, E. A., Danforth, C. M., and Dodds, P. S. (2015). Characterizing the Google Books corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PLoS ONE*, 10(10):e0137041.
- Sageman, M. (2017). *Turning to Political Violence: The Emergence of Terrorism*. University of Pennsylvania Press, Philadelphia.
- Tahmasebi, N. (2018). A study on Word2Vec on a historical Swedish newspaper corpus. In *Proceedings of DHN 2018*, pages 25–37, Aachen. CEUR-ws.org.
- Tahmasebi, N., Borin, L., Capannini, G., Dubhashi, D., Exner, P., Forsberg, M., Gossen, G., Johansson, F., Johansson, R., Kågebäck, M., Mogren, O., Nugues, P., and Risse, T. (2015). Visions and open challenges for a knowledge-based culturomics. *International Journal on Digital Libraries*, 15(2–4):169–187.
- Tahmasebi, N., Dannélls, D., and Adesam, Y. (2019a). Exploring the quality of the digital historical newspaper archive KubHist. In *Proceedings of DHN2019*, pages 9–17, Aachen. CEUR-ws.org.
- Tahmasebi, N., Hagen, N., Brodén, D., and Malm, M. (2019b). A convergence of methodologies: Notes on data-intensive humanities research. In *Proceedings of DHN2019*, pages 437–449, Aachen. CEUR-ws.org.
- Verhoeven, C. (2009). *The Odd Man Karakozov: Imperial Russia, Modernity, and the Birth of Terrorism*. Cornell University Press, Ithaca.
- Viklund, J. and Borin, L. (2016). How can big data help us study rhetorical history? In *Selected Papers from the CLARIN Annual Conference 2015*, pages 79–93, Linköping. LiUEP.
- Zimmer, B. (2013). When physicists do linguistics. Is English ‘cooling’? A scientific paper gets the cold shoulder. *Boston Globe*, February 10.