HAL open science

Toponyms as Entry Points into a Digital Edition: Mapping Die Fackel (1899-1936)

Adrien Barbaresi

▶ To cite this version:

Adrien Barbaresi. Toponyms as Entry Points into a Digital Edition: Mapping Die Fackel (1899-1936). Digital Humanities, Aug 2017, Montréal, Canada. pp.159-161. hal-01591628

HAL Id: hal-01591628

https://hal.science/hal-01591628v1

Submitted on 21 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Toponyms as Entry Points into a Digital Edition: Mapping Die Fackel (1899–1936)

Adrien Barbaresi
adrien.barbaresi@oeaw.ac.at
Austrian Academy of Sciences, Austria

Introduction

The significance of place names exceeds the usually admitted frame of deictic and indexical functions, as they enfold more than a mere reference in space. In the western tradition, a current of reflexion which seems to date back to the 1960s has provided the theoretical foundations of the "spatial turn", whose epitome is the concept of space as emergent rather than existing a priori, and composed of relations rather than structures (Warf, 2009). The emergence of current named "GeoHumanities" (Dear et al., 2011) or "Spatial Humanities" (Bodenhammer et al., 2010), has prompted for a transfer of research objects between disciplines as well as an enforcement of the spatial turn in practice through specific methods of analysis. The common denominator consists in opening up new spaces and experimenting in a transdisciplinary perspective (Domínguez, 2011) in a field which has been evolving at an exponential pace since the last decade (Caquard and Cartwright, 2014).

In this paper, I introduce a visualization of collocations of toponyms in the satirical literary magazine Die Fackel ("The Torch"), originally published and almost entirely written by the satirist and language critic Karl Kraus in Vienna from 1899 to 1936. This work carries heterogeneity at its core and contains a considerable variety of toponyms (Biber, 2001) which are highly significant because of the multinational nature of the Austro-Hungarian empire and the later formation of a territorially diminished state.

In order to provide an additional, synthetic access to a digital edition of the work which is already available online (AAC-Fackel corpus), I set out on a distant reading experiment leading to maps meant to uncover patterns and specificities which are not easily retraceable during close reading. I focus on the concept of visualization, that is on the processes and not on the products (Crampton, 2001), and present them together with a critical apparatus, by giving a theoretical perspective on what is being shown and seen. In fact, digital methods in humanities ought to be criticized (Wulfman, 2014) and the cartographic enterprise bears both a thrill and a risk: "adding more to the world through abstraction", and "adding to the riskiness of cartographic politics by proliferating yet more renders of the world" (Gerlach, 2014).

Extraction of toponyms

The particular task of finding place names in texts is commonly named place names extraction, toponym resolution, or geocoding. A first stage involves the identification of potential geographic references, while a second stage resides in a disambiguation process (Leetaru, 2012). Toponym resolution often relies on named-entity recognition and artificial intelligence (Leidner and Lieberman, 2011). However, knowledge-based methods using fine-grained data – for example from Wikipedia – have already been used with encouraging results (Hu et al., 2014).

The present endeavor grounds on a specially curated gazetteer: during the 20th century there have been significant political changes in Central Europe that have severely affected toponyms, so that geographical databases lack coverage and detail. Consequently, the database developed at the Austrian Academy of Sciences (Academy Corpora) in cooperation with the Berlin-Brandenburg Academy of Sciences (Language Center) focuses on Europe and follows from a combination of approaches: gazetteers are curated in a semi-supervised way to account for historical differences, and current geographical information is used as a fallback. Wikidata API and the Geonames database are used to build the database semi-automatically.

The tokenized files of works to be analyzed are filtered and matched with the database by finite-state automatons (Barbaresi and Biber, 2016): toponyms (single or multi-word expressions) are extracted using a sliding window. A cascade of filters is used: current and historical states; regions, important subparts of states, and regional landscapes; populated places; and geographical features. Disambiguation being a critical component (Leetaru, 2012), an algorithm similar to Pouliquen et al. (2006), who demonstrated that an acceptable precision can be reached that way, guesses the most probable entry based on distance to Vienna (Sinnott, 1984), contextual information (closest-country, last names resolved), and importance (place

type, population count). The results are projected on a map of Europe using TileMill.

From collocations to lines of thought

In a further analysis, I visualize co-occurrences of extracted toponyms, which can be considered to be a subset of GeoCollocations (Bubenhofer, 2014), in order to draw sequences, airborne lines following their order of appearance. The word "network" is to be used with circumspection as Latour (1999) suggests. Although it is ubiquitous in the terminology of the spatial turn, the now predominant interpretation in the sense of the World Wide Web suggests an immediacy which is contrary to the acceptions it had before, so that the concept of "meshwork" is more appropriate (Ingold, 2007). I thus interpret Figure 1 as a general meshwork which makes it possible to visualize paths depicting chains of thought (Gedankengänge) as well as their intensity (well-trodden or seldom). If they may reveal spatial patterns that would otherwise remain hidden in texts (Bodenhammer et al., 2010), these linkages are also "mappings and tracing imposed on the data" (Wulfman, 2014) which are not meant to be interpreted without further filtering.

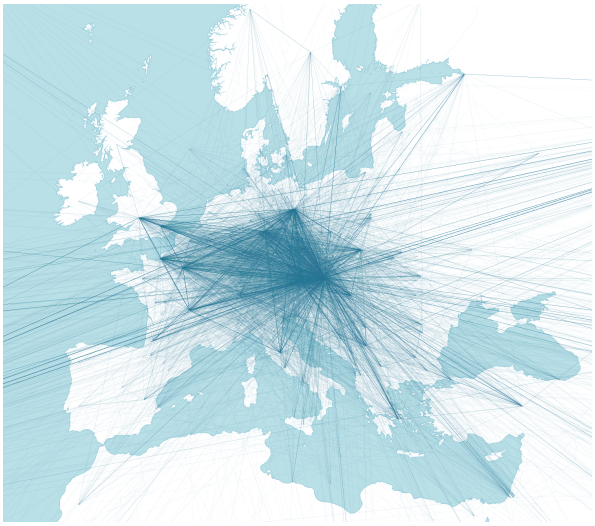

Figure 1. Unfiltered map of toponymic co-occurrences

A rhizome as entry to Die Fackel

That is why I refine the map by selecting a subset of the co-occurrences – the maximal distance between two extracted place names is fixed to twenty tokens – and by color-coding qualitative features – the typology of place names and the axis of time. The most frequent place names are printed out. Surfaces (regions for instance) cannot be represented as such because of historical evolutions and because of the difficulties of linking surfaces without tampering with map readability. Coastlines are depicted in gray to give a sense of orientation, no boundaries are drawn, as they are of a changing nature and may superimpose an artificial reading of the map (Smith 2005).

Figure 2. Refined analysis (size proportional to corpus frequency; yellow: sovereign territories; orange: regions; green: populated places; blue: geographical features; time axis represented by a gradient from light green to dark blue)

The notion of rhizome has been used in corpus linguistics by Scharloth et al. (2013) to qualify discourses captured by collocation graphs, it has originally been coined by Deleuze and Guattari (1987 [1980]). This concept is particularly adequate for Kraus, as the Austrian satirist has always been concerned by the multiple aspects of discourse and reality. In addition, his work in Die Fackel evades distant reading processes because of the number of citations used and its ever present and extensive usage of parody. It would be vain to design an authoritative cartography of Die Fackel: following from the principles of heterogeneity and "asignifying rupture" (ibid.), the lines are frequently interrupted. Phenomena in the low-frequency range are filtered out by the human eye, but clusters and interpretation cues may emerge which provide a different access to the work. In this regard, Figure 2 depicts a rhizome connecting heterogeneous information, just as we are all "traversed by lines, geodesics, tropics, and zones marching to different beats and differing in nature" (ibid.).

Conclusion

I have presented a distant reading experiment which consists of connecting toponyms extracted and

projected on maps. The latter are meant to be released as an additional feature to the existing digital edition. The Software and gazetteer will be made available under open-source licenses, for development files (please refer to the Github repository). The first example displays unfiltered lines of thought, whereas the second one grounds on a refined analysis and lets the practical image of a rhizome emerge. While Die Fackel criticizes mechanical, instrumental language (Hirt, 2002), the "well-informed" linguistic instruments can help materializing dots or sequences, but not without "human" intervention. The filtering steps on the projection echo the hermeneutic circle of the philological tradition; they also make computed information visible and apprehensible which could remain "blind" otherwise (Barbaresi, 2012).

This is not an authoritative cartography of Die Fackel but rather an indirect depiction of the viewpoint of Kraus and his contemporaries. Drawing on Kraus' vitriolic recording of political life, toponyms in Die Fackel tell a story about the ongoing reconfiguration of Europe. As the map conveys the uncanny sensation that the continent is a field on which points east and west are projected, the lines of force entangle European countries and capitals. Their spatial patterns document an inclination for major cultural centers, whereas the chronological dimension captures a major shift towards the end of publication: the force field intensifies as its range narrows, showing both the interplay of major European powers of the time and the emergence of transatlantic (westwards) and transeuropean (eastwards) relationships. This evolution can be read as an intensification of tensions and a prefiguration of other schemes, this time of military nature.

Bibliography

AAC – Austrian Academy Corpus (2007). *AAC–FACKEL, Online Version: "Die Fackel. Herausgeber: Karl Kraus, Wien 1899-1936".* http://www.aac.ac.at/fackel

Barbaresi, A. (2012). "La Raison aveugle ? L'époque cybernétique et ses dispositifs". Conference manuscript: *Les critiques de la raison au XXe siècle*, University Paris-Est-Créteil, 2012.

Barbaresi, A. and Biber, H. (2016). "Extraction and Visualization of Toponyms in Diachronic Text Corpora". *Digital Humanities 2016: Conference Abstracts*, pp. 732–734.

Biber, H. (2001). "In Wien, in Prag und infolgedessen in Berlin – Ortskonstellationen in der 'Fackel'". Marten-Finnis, S. and Uecker, M. (eds), *Berlin-Wien-Prag. Moderne, Minderheiten und Migration in der Zwischenkriegszeit*, Peter Lang, pp. 15–26.

Bodenhamer, D. J., Corrigan, J. and Harris, T. M. (2010). *The Spatial Humanities: GIS and the Future of Humanities Scholarship*. Indiana University Press.

Bubenhofer, N. (2014). "Geokollokationen – Diskurse zu Orten: Visuelle Korpusanalyse". *Mitteilungen des Deutschen Germanistenverbandes*, 61(1), pp. 45–89.

Caquard, S. and Cartwright, W. (2014). ""Narrative Cartography: From Mapping Stories to the Narrative of Maps and Mapping". *The Cartographic Journal*, 51(2), pp. 101–106.

Crampton, J. W. (2001). "Maps as social constructions: power, communication and visualization". *Progress in Human Geography*, 25(2):235–252.

Domínguez, C. (2011). "Literary Geography and Comparative Literature". *CLCWeb: Comparative Literature and Culture*, 13(5), 3.

Dear, M. et al. (2011). *GeoHumanities: Art, History, Text at the Edge of Place*. Routledge.

Deleuze, G. and Guattari, F. (1980). *Mille Plateaux.* Éditions de Minuit. English translation: 1987, University of Minnesota Press, translation by Brian Massumi.

Foucault, M. (1984). "Of Other Spaces, Heterotopias". *Architecture, Mouvement, Continuité*, 5, 46–49. Original Publication: *Conférence au Cercle d'études architecturales*, March 14th, 1967.

Gerlach, J. (2014). "Lines, contours and legends. Coordinates for vernacular mapping". *Progress in Human Geography*, 38(1):22-39.

Hirt, A. (2002). *L'Universel reportage et sa magie noire. Karl Kraus, le journal et la philosophie.* Kimé.

Hu, Y., Janowicz, K. and Prasad, S. (2014). "Improving Wikipedia-Based Place Name Disambiguation in Short Texts Using Structured Data from Dbpedia". *Proceedings of the 8th Workshop on Geographic Information Retrieval*, ACM, pp. 8–16.

Ingold, T. (2007). *Lines: A Brief History*. Routledge.

Latour, B. (1999). "On recalling ANT". *The Sociological Review*, 47(S1):15–25.

Leetaru, K. H. (2012). "Fulltext geocoding versus spatial metadata for large text archives: Towards a geographically enriched Wikipedia". *D-Lib Magazine*, 18(9), 5.

Leidner, J. L. and Lieberman, M. D. (2011). "Detecting Geographical References in the Form of Place Names and Associated Spatial Natural Language", *SIGSPATIAL Special*, 3(2):5–11.

Smith, M. L. (2005). "Networks, territories, and the cartography of ancient states". *Annals of the Association of American Geographers*, 95(4), pp. 832-849.

Scharloth, J., Eugster, D. and Bubenhofer, N. (2013). "Das Wuchern der Rhizome. Linguistische Diskursanalyse und Data-driven Turn". *Linguistische Diskursanalyse: neue Perspektiven*. Springer Fachmedien Wiesbaden, pp. 345–380.

Warf, B. and Arias, S. (2009). "Introduction: the reinsertion of space into the social sciences and humanities". *The Spatial Turn: Interdisciplinary Perspectives.* Routledge, London.

Wulfman, C. E. (2014). "The Plot of the Plot: Graphs and Visualizations". *The Journal of Modern Periodical Studies*, 5(1):94–109.