Symbolic dimensions of serial sources: hermeneutical problems of reconstructing political biographies based on computerized record linkage

Lipp, Carola

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY-NC Lizenz (Namensnennung-Nicht-kommerziell) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: https://creativecommons.org/licenses/by-nc/4.0/deed.de

Terms of use:
This document is made available under a CC BY-NC Licence (Attribution-NonCommercial). For more Information see: https://creativecommons.org/licenses/by-nc/4.0

Diese Version ist zitierbar unter / This version is citable under:
http://nbn-resolving.de/urn:nbn:de:0168-ssoar-34268
Symbolic Dimensions of Serial Sources.
Hermeneutical Problems of Reconstructing
Political Biographies Based on
Computerized Record Linkage.

Carola Lipp *

Abstract: This article presents a project, which seeks to reconstruct the political culture in a German town (Esslingen) from 1830 to the Revolution of 1848. It is based on computerized record linkage and connects social and familial data from serial sources with qualitative information about political attitudes and behavior. The core of this prosopographical approach is produced by individual and collective biographies that show the meaning of political behavior within the lifecycle and lifestyle of people of different social classes. In addition, linking these biographies reveals social and kinship networks that underlie political activities. The article deals especially with methodological problems of identification and individual self-representation in serial sources that complicate nominal record linkage procedures. In petitions for example people present themselves in different roles and with different interests. Therefore a simple list of names can give unexpected insights into social relations and the self-interpretation of people, groups and organizations.

The contemporary debate in sociology and social history has been dominated by the dualism between quantitative approaches based on statistical procedures and taxonomic theories about social organization and development and a qualitative interpretive approach which focuses on the individual and the meaning of his/her action (1). In actual historical research these paradigmatic boundaries are not as strict as has been postulated. Anthropological approaches in history, especially local and demographic microstudies, try to bring both approaches together and to come to a micro-macro-link (2). The research project I will present is com-

* Address all communications to Carola Lipp, Universität Göttingen, Institut für Volkskunde, Goßler Str. 5-7, D-3400 Göttingen.
mined to the new cultural history as well as to the tradition and techniques of social history. It deals with the revolution of 1848 and is especially interested in the connection between political culture and everyday life (3). Political culture means in this context not only organizations, political institutions or direct political action, but means also the vast complex of values, attitudes and behaviour patterns (4) that influence political decisions, structure political life as a world of social interaction and create meaning. So far studies on gender participation, associations and public rituals in riots or festivities have been completed (5), but the center of the project, is a large database that includes information about action people. It is based on nominal record linkage of serial and textual sources. This article will deal with questions and problems resulting from the procedure of record linkage. I want to show that even serial sources have self-interpretive dimensions and that this fact affects linkage decisions and the operational range of analysis.

The object of this research project is the entire population of the city of Esslingen in Wuerttemberg and their political activities leading up to and including the revolution of 1848. It is intended to reconstruct individual and collective political biographies in order to build the core of a complex analysis of the political culture at that time. Esslingen was a very busy place in 1848; there were many political declarations and active organizations. The sources are rich. Especially the newspaper is filled with political pronouncements of various groups, which allows one to identify changing political factions. The complete data base includes three types of sources: a social data base, demographic data from church records, and political sources (see graphic). The latter include files of different types of information, among others, a large series of petitions, information about positions in the city council and local organizations, membership rolls of associations, data from court records and the city council protocols, and detailed extracts from the local newspaper concerning single activities of individuals and groups. The local electoral lists form a data base in itself, which delivers not only names and professions, but the individual choice of each voter for a period from 1840 to 1848.

Of central interest are different styles of political behavior and modes of articulation and action related to certain groups or individuals. [The multiplicity of sources shows differences in cultural practice and in the appropriation of cultural techniques. It is interesting to know which groups use cultural media like the newspaper as a forum for political announcements, or tend to sign petitions, or like to riot at night. Often the group boundaries are considerably more flexible than cultural prejudice assumes. The time certain people enter the political scene, or join organizations, sign petitions or demonstrate publicly, is as important for the meaning of the action as the question of how it fits into their lifecycle.
How are their activities related to the group they join, or live or work within? Which bounds do they share? Are there processes of political adaptation? How and in which context do people learn certain styles of political behavior, etc.? These are some of the questions that are posed. The result will be a »thick description« (6) of specific actions and groups, using descriptive and analytical statistical procedures as well as hermeneutical interpretation of the symbolic dimensions of political action and of the meaning contemporaries have inscribed. Conditions of life experience, social relations and self-perception shape political behavior and thinking, just as political experience forms and changes everyday life. Political culture is understood as embedded in structures of life experience and habitus (7). To reveal social and private networks (8) underlying political activities and thus to get an idea of the inherent structures of meaning is one of the major goals of this project. These networks include *neighborhood*, which emerges as a force strong enough to transcend social stratum, and *workplace*, both structures with a heavy impact on political behavior. They also include *family relations*. The relationship between class formation processes, political organization and strategies of social and individual reproduction, like marriage, will be one of the important fields of analysis. Scholars often repeat the phrase that organizations are marriage markets (9), but no one could ever show it. By the use of demographic data, it is not only possible to link political behavior to structures of individual life, but also to analyze demographic behavior from a political point of view. The problem of *generation*, for example, is a central one in 1848, with regard to both intergenerational conflicts and the persistence of traditional models of political behavior.

Political attitude in this regard means not only public activity, but also behavior in a domain which is still called the private sphere. One might ask, for instance, if self-identified democrats pursued marriage patterns (10) consonant with their politics. This study reaches far beneath the level at which studies of political culture usually stop. If, for example, the liberal party tried to survive in the forties by organizing children's festivals and claiming to be the party of the future, the question might be asked if and how this idea showed up privately in certain family patterns and if there was a consistency between political and private goals.

Contrary to other studies, the project does not concentrate on prominent groups or restrict its view to the social structures of certain politically dominant organizations (11). Instead I try to look at the whole system of organizations and cross-memberships to determine how an activity is related to other forms of political participation, like signing petitions, preferences in voting, membership in cultural organizations, positions in the local executive, and so on. What makes this study especially interesting is its focus on the city as a whole. One can thus not only identify the political
activists, but also the non-participants. It is as important to learn, among those permitted to vote for the National Assembly, who did not vote as to learn who did. Which are the structures that explain non-participation? Does the answer lie in socioeconomic constraints or, as is in fact the case, in the neighborhood and in the self-definition of certain professions (like judges and priests)? Political attitude is thus explained in the context of an entire lifestyle, a whole construct of cultural activity. One part of the study thus deals with processes of political formation and organization. The other part concentrates on social substructures of political life.

Methodologically, the concept is based on several steps of nominal record linkage, that integrate data bases of different information levels and logical structures. The file management as well as the record linkage are done by KLEIO, a data bank oriented system (12) that handles dataformats quite flexibly. KLEIO provides differentiated procedures to create a relational net between separate files; it is even able to merge files of different structure physically and thus combine the information of several sources in to a new file. Nine files have been linked. At the end nearly forty larger and smaller sized files will be matched. Naturally, the increase of information with each linkage quickens further identification routines.

The starting point of the project was a data base containing central data of the population of Esslingen. This social data base is founded on two inhabitant lists which provide basic information about the heads of more or less 2,250-2,500 households. Besides the names, last and first names, sex and dwellings, the population registers include information about the profession and sometimes even the place of work. Added to these files is an inhabitant-tax roll which catches nonresident people, the mobile group of inhabitants who only stayed for some months. Electoral lists for the national assembly and militia records with detailed information about the actual occupation and birth date and a list of house owners complete this complex, which consists of approximately 11,000 nominal entries, representing a population of 12,763 inhabitants. These data relate to approximately 3,600 adults, mostly male, because the female population is represented only by widows and self-supporting working women. Referring to a period of 6 years from 1846 to 1852, the net of social data is thus tightly knit. All the information, linked and non-linked entries, out of the different files is integrated in one complex social data base without changing the shape of the single file or discarding information. Non-matching cases are not dropped, but live on as »singles« in the central data base with an identification number that shows the file they originate from. Often these »singles« find their match in another linkage step. The database therefore is larger than the population in reality. The percentage of linkages differs from 55% to 95% depending on the character of the source and the items contained (13).
Technically, the linkage procedure is done semi-automatically; that means the computer assists in preparing decisions, but does not execute the linkage. System routines in the form of repertory lists (14) support the detailed comparisons of records. The choice is limited by the definition of selection criteria like surnames and forenames, occupation, age, street and number of the house. The system starts with an exact correspondence of strings and then relaxes the set of linkage variables in subsequent steps. The identification is controlled by the researcher, who also makes the final linkage decision and the definition for the matching of two files. Though the computer assistance supports identification, there are cases where known people are more easily identified and quicker linked by hand than by machine, because of the simplicity of the selection process. This is especially true for small files (15). Defining too many variables lames the computer and ends in a low number of proposals; defining a broad frame produces too much paper and too many alternatives. In complicated cases the machine and KLEIO's comparing functions are indispensable. KLEIO and its integrated Soundex algorithm offer the possibility of overcoming problems of misspelling, spelling variations, homonyms, or other minor inaccuracies, which usually cannot be recognized by manual inspection of the file. The Soundex algorithm reduces alphabetical strings to significant and differentiating combinations and allows one to define variations of word endings and beginnings, a procedure which enhances the number of identifications enormously.

In a further step the linkage of social data is controlled and broadened by information from church records. These are Protestant, Catholic and Jewish family registers that include three generations, the marrying couples and their parents, as well as the children and the spouses of the second generation. Since the church records cover a period from 1800 to 1870, these sources allow one to reconstruct the intergenerational profile of the residential population. Both linkage processes, the family reconstitution as well as the social data base, are made parallel and used as a controlling list of items for the linkage process. After the completion of the internal linkage of the family records, this source will be linked to the central data base. Thus the relational net of the family will be tied together with political, social and topographical data. The family reconstitution is not yet finished, but the social data base (as it is shown on the graphic) is nearly accomplished and partly linked with data from membership rolls and petitions.

Besides the central problem of record linkage, the question if it is a false or »true« link, record linkage always creates a social reality in itself, »a theoretical construct« and therefore represents a complex hermeneutical act, a »calculation of possibilities« (16). Besides this there emerge other more practical problems with this approach. The technique of integrating a
complex number of closely related data into one single file leads to multiple entries which are difficult to use in statistical analysis. The social data base, for example, shows more than 6 profession entries from different periods and each linkage might differentiate the information and give not only the profession, but also the workplace and the exact status. Thus the question arises: Which occupational definition is to be used for statistical purposes, when I examine a linked source which originally does not contain information about the occupation? The structure of the central file therefore has to be flexible to reflect these modifications.

Furthermore, the information in the entries usually varies quite remarkably. In a published official source like the printed inhabitant registers, an artisan might be keen to show up in his learned profession, though the militia list reports that he worked at that time in a factory. How is he to be treated? On the one hand, the register reflects a traditional status which could be the sign of his clinging to this social position. And this tendency might also be observed in a formal letter to the city council. On the other hand, it is important to file him as a factory worker when he has signed a petition that runs through the staff of a certain factory. To subsume such a person simply under the uniform category »actual profession« excludes an important social dimension. Obviously he himself wanted to be presented in a different way.

Methodologically, this problem could be solved by extrapolating certain career steps out of the material. Thus, ambiguous information can be structured along a longitudinal axis. Defining special thesauri for such groups of transitional professions is another way to handle this problem. And the last method would be to separate analytical procedures. Therefore each source has to be analyzed at first in its own given social horizon and a second time with the aggregated career definition, which is developed from the linked information and then mapped on to single files. As the career definition gets more specified with the accumulation of data, the mapping of this information to other files produces interesting results. For the analysis of population structure or acting people, this means that sometimes a quite different social reality shows up. Besides the general danger that ambiguities or differentiation might be produced by false identifications and linkages, my data base reveals certain characteristics of serial sources which are usually ignored or cannot be detected. Printed inhabitant lists at that time, for example, tend to raise social status or to preserve a certain social profile (17). Factory working is noted as a profession only in cases where factory working is regarded as something special.

Between 1847 and 1850 Esslingen underwent rapid social change caused by the founding of one of the most important machine factories in the south-west. (18) Population changed quickly at that time and so did the professional profile of the population. In the register of 1850, the skilled
workers of the Esslinger Maschinenfabrik emphasize this fact in reporting their workplace, whereas workers in the textile factory use their learned profession and do not refer to their actual status. The question remains, whose value system is reflected by inhabitant registers? I think there is a tendency toward respectability which reflects the self-reference of the documented people and which has its consequences not only in the representation of lower classes, but still more in the representation of upper classes. Well-to-do owners of several houses or business places appear more than once in the registers without indication of their double identity. Cross checks with church records, therefore, show that the upper classes are usually overrepresented. This fact might be observed in other towns as well. It is a known result of record linkage not only to produce faults but to detect errors und inaccurate entries in the original document. (19)

Therefore, the constituency of sources is very important. People, especially high ranking people, play a vast number of social and political roles and this influences their representation in certain sources. I was confronted with this fact in linking my social data base with petitions. Usually the petitions I am dealing with include only name and surname and, under certain conditions, they tell the profession. And there the problems of identification start. A petition addressed to the National Assembly in Frankfurt pleading for heavy duties on the import of wine was signed only by vintners and vineyard-owners (20). The linkage to the social data base revealed that the latter category contained a large number of non-agrarian professions like coppersmiths, butchers or tanners, who referred to themselves as vineyard-owners, because this was reflecting their actual interest. One of these people, a confectioner, appeared in a different petition as merchant (Handlungsvorstand); and, indeed, he owned a little factory that produced vinegar, a fact that certainly gives a bad reputation to his vineyard. That the same man also liked to use his titles of Gemeinderat or commander of the civil army is not astonishing. His self-referential notations changed, adapted to the situations and groups he was acting in. Traditional statistical analyses of serial sources, especially petitions (21), usually do not reflect this interactive dimension of sources and change of social roles (22).

Though there is a logical structure in this behavior, these conditions enormously complicate record linkage of process-produced sources. However, it is this inherent social code which can sometimes lead to the identification of people. There exist certain rules of self-perception and contextual knowledge that guide signing petitions. If the leader of a petition or the one who signs first on a page uses a professional title, the others will follow. Mostly petitions do not report professions because they are produced in an already defined social group. The existence of occupational remarks in a petition therefore often indicates that there is a political, but
no obvious social or situative, relation between the acting people. If the signing people are aware of the common patterns or structures they are acting in, they tend not to give contextual information. When occupational connotations appear in these cases, they are meant to demonstrate social difference and thus break a logical sequence in the petition. In a list of 20 factory workers who do not sign with their profession, the learned baker might insist on his different background. When a petition passes a street, for example in a vintner’s quarter where closely interrelated people live, people who have the same family names and often the same profession usually try to identify themselves socially by writing the name of their father, or surnames. Very often the last one in the row omits these connotations because he might have thought that after all the others had signed he is socially placed and needs no further identification. The same rule is followed by a council member who came from a big vintner family and who never signed with his first name or his profession because he obviously expects to be known. And at last, after much puzzling, I can identify his rather prominent signature. When I am talking about symbolic dimensions of serial sources, I mean these forms of self-presentation and self-reference that should be reflected in record linkage processes.

Process-produced data that emerge out of context-defined political situations like an association or neighborhood, usually contain very little hints to identify persons. Therefore, a large social data base and a systematic testing of social networks is required if those data are to be linked. When this procedure is successful, record linkage is then able to reveal social structures that cannot be seen at first glance. A last example: Petitions addressed to the National Assembly usually carry titles which are given to them by the registration bureau. An Esslinger petition pleading for protective duty (23) was thus assigned to merchants and factory owners who headed the list. Closer inspection after record linkage, however, showed that nearly 60% of the people who had signed were factory workers, a fact which enormously affects the interpretation of the material. Used in such a way, record linkage, for example, allows one to follow the social route a petition takes. It can be shown in which groups, neighborhoods or workshops the lists had circulated.

Complex record linkage thus reveals the underlying social network and mechanisms of political actions; the integration of single files with a central database allows one furthermore, to look at interrelated cultural patterns. The strength of this approach lies in its flexibility. It can focus on individuals and groups as well, and analyze common structures and rules of action. It is possible to examine the social context of a whole organization, as well as to concentrate on certain prominent figures in the local elite. One can survey political activities and also the private lives of the actors. Multilevel analyses and testing make it easy to look for dominant
features of political behavior. KLEIO’s capacity to manage complex file structures, the flexible and rich retrieval function support not only record linkage processes, but make it possible to decode symbolic and self-referential dimensions of serial sources and thus enrich interpretive possibilities. Canned routines which translate complex file structures into an arbitrary number of SPSS-matrices open the way to descriptive statistics as well as complex multivariate analyses. But the study does not end with statistical analyses, procedures like cluster or dominance analyses shall help to identify groups with common patterns whose actions will then be analyzed in their individual (as well as collective) biographical and social context. Thus statistics opens the way to the individual and his/her way to participate in the political culture 1848/49.

Notes


(3) This project was supported by the Deutsche Forschungsgemeinschaft. An outline of the original research plan can be found in: Wolfgang Kaschuba, Carola Lipp, »Revolutionskultur 1848/1849«, in: Helmut Reinalter (ed.), Aufklärung-Vormärz-Revolution, vol. 4 (1984), pp. 20-24.


(5) Carola Lipp (ed.): Schimpfende Weiber und patriotische Jungfrauen. Frauen im Vormärz und in der Revolution 1848 (Bühl-Moos, 1986). Wolfgang Kaschuba, »Vom Gesellenkampf zum sozialen Protest«, in: Ulrich Engelhardt (ed.), Handwerker in der Industrialisierung (Stutt-


(10) Karin Hausen, for example, showed that the large age difference in bourgeois marriages produced an assymetrical relation between husband and wife and put the wife into the social position of a grown up child. K. Hausen, »Ulme für das schwanke Efeu«. Ehepaare im Bildungsbürgertum. Ideale und Wirklichkeiten im späten 18. und 19. Jahrhundert,« in: Ute Frevert (ed.), Bürgerinnen und Bürger (Göttingen 1988), pp. 17-48.

(11) See studies regarding the development of the party system, the liberal and democratic movement and working class organizations. Werner Boldt, Die Anfänge des deutschen Parteiwesens (Paderborn, 1971); Dieter Langewiesche, Liberalismus und Demokratie in Württemberg (Düsseldorf, 1974); Peter H. Noyes, Working Class Associations in German Revolutions of 1848-1849 (Princeton, 1966); Frolinde Baiser, Sozial-Demokratie 1848-1863, 2 vols. (Stuttgart, 1962) or Wolfgang Häusler, Von der Massenarmut zur Arbeiterbewegung (München, 1979).

(13) The inhabitant lists, for example, refer to independent households or to heads of households. The militia list includes each man in the city between the age of 20 and 50. The first linkage rate therefore is only 55% on the first level of record linkage. After the addition of other files this rate grows naturally. The percentage of identifications between two militia lists lies at 95%.


(20) Bundesarchiv Frankfurt, Bestand DB 51, Bü 213.

(21) Heinrich Best, Interessenpolitik und nationale Integration. Handels-politische Konflikte im frühindustriellen Deutschland (Göttingen, 1980).


(23) Bundesarchiv Frankfurt DB 51, Bü 218b.