

A multi-level database : study on the elite of 15th and 16th-century Groningen.¹

REDMER ALMA

1. Introduction

This contribution reports on a long-term study of the elite groups in the city and province of Groningen (formerly the Frisian 'Ommelanden') during the late Middle Ages. The situation with regard to the sources for prosopographical research in this area differs somewhat from that of other regions in the Netherlands. Moreover, the structure of the elite groups studied does not entirely correspond to the general pattern. Both of these factors have influenced the organization of the discussed database. Therefore, I will begin by discussing the sources and some of the idiosyncrasies of the upper classes in both the city and province of Groningen. The main body of this article is devoted to the difficulties involved in setting up a database, and is followed by a discussion of the multi-level structure of this database, as well as a number of general conclusions.

2. Sources

The city of Groningen and the *Ommelanden* to the north of it seem an unlikely subject for a prosopographical study focusing on the late Middle Ages. For one thing, the archival sources on the city of Groningen is rather scarce as compared to the other cities in the Netherlands : municipal accounts, fiscal rolls and judicial or administrative protocols before 1550 are almost entirely absent and lists of councillors -essential to the research of urban elites- do not exist at all. The absence of such material is not due to fire, floodings or other natural disasters. Rather, it is the result of the administrative situation in Groningen at that time. Obviously, valuable documents such as mortgage deeds and charters were stored, but documents that were not of permanent value to the city were often kept by the civil servants who dealt with them. Urban accounts were simply returned to the treasurer after clearance. After repeated unrest in the city, a professional municipal administration was finally set up, together with the necessary creation and administration of archives in the mid-sixteenth century, under pressure of the successive princes. In the *Ommelanden*, such a system did not exist, which is hardly surprising, given

¹ Thanks are due to E.W. van Royen for reading through the article and making many valuable comments and to Barbara Corrick for correcting the translation.

that it was not until the sixteenth century, that the region's administrative bodies were unified.

Hence, the inevitable question: what sources are actually available? Fortunately, we do have more or less comprehensive lists of mayors and *hoodmannen* (judges of the court of civil justice) compiled by Ubbo Emmius. In addition, there is a number of protocols, taxation lists drawn up in the fifteenth and sixteenth century and the so-called '*clauwboeken*', manuscripts, which record the possession of the yearly rotating right to appoint the local judge. The 20,000 odd records from the city's archives, churches, monasteries and other institutions also comprise an important part of the available archival sources. In addition, private archives, such as home, estate and personal archives contribute significantly to the total number of documents. All these sources provide information primarily on the transfer of real estate and legal titles.

The quality of the archival sources handed down, differs greatly from one archive fund to the other, both with respect to distribution over time and place, as well as to the individuals about whom information is conveyed. This puts great restrictions on the methodology of the researcher. As a result, it is virtually impossible to quantify data on prosopographically important questions. This problem is not faced by prosopographers alone. Whether or not well-documented families such as the Lewe, the van Ewsum and the Rengers deserve such a prominent place in the historiography of Groningen should also be a point of debate in traditional historiography. There is little documentation on families as the Onsta, the Cater or the Tamminga, though their only crime was not to leave behind an estate or a family archive.

Due to the lack of coherent archival sources, superficial prosopographical research might lead us to the same results that could be drawn from the existing literature: a list of well-known names of rich and powerful individuals all related to one another and a number of *homines novi* ready to fill vacant positions or to move up in any other way. The task of checking, and if necessary adjusting, the traditional image of social and geographical mobility, requires closer study of the sources available. It is especially the combination of a varied source material that yields much new information. Yet, we must never lose sight of the fact that much of the archival sources are incomplete.

3. Problems

The literature on the elite in the province of Groningen, paints a picture full of anachronisms. The 15th- and 16th-century ancestors of 17th and 18th-century nobility are given in retrospect the same high social status

as their descendants. Little attention is paid to *hoofdelingen* (gentry) and patricians who did not bear a heritable name, but merely a patronym or toponym. However, the structure of the nobility in the *Ommelanden* during the late Middle Ages appears to differ greatly from that of the 17th and 18th centuries: one's status as a *hoofdeling* in those days was determined by succession in the paternal line as well as in the maternal line and by primogeniture. Succession of the *hoofdeling* status is never mentioned explicitly in contemporary sources: it cannot be traced in division of property, wills, legal documents, etc. Yet, the marital partners of *Hoofdelingen* are predominantly chosen out of the children of other *hoofdelingen*. At first glance, this seems hardly surprising, but that consistent a policy on marital partners is striking in such a small group of which the cohesion is formed by nothing more than a concept of nobility based on unwritten rules.

We have focused here on the *hoofdelingen*, not so much in order to present our findings, but to illustrate the extent to which their structure has influenced our methodology. Because the factors determining the status as *hoofdeling* are not yet entirely clear and the number of *hoofdelingen* is relatively small, we did not consider it worthwhile to conduct a sample survey of the *hoofdelingen* and make a closer study of the individuals encountered in that group in order to draw general conclusions. It is difficult to take a representative sample from a small and ill-defined group of people as the chances are considerable of including an atypical individual. Moreover, there are some questions that require information on the entire population. For instance, we stated above, that the connubium (in respect to the marital partners from the *Ommelanden*) was significant. The question whether the class of *hoofdelingen* was indeed completely closed with respect to regional intermarriage, would disappear from view in a sample survey. The same applies to the question whether all *hoofdelingen* owned rights to appoint judges or lived in a 'stonehouse' (*steenhuus*).

Although the task of establishing a manageable, well-defined group of the elite seems feasible for the *Ommelanden* region, it entails far more difficulties for the city of Groningen itself, a factor which also influences the methodology. The lack of sources on the occupation of municipal positions thwarts research. There is no choice but to abandon attempts to separate the higher social classes from the lower ones in the period before 1520. Nevertheless, it is possible to identify individuals who belonged to the political elite of 15th century Groningen. Among these individuals (mayors, chiefs (*hoofdmannen*) and councillors), a group emerges that differs with respect to family relations and political careers from those who presumably did not belong to the urban political elite. Moreover, within the elite, subgroups

emerge of families which rank higher or lower on the social scale. It is still unclear to what extent the social, political and economic elite in the city of Groningen coincided.

4. Effects on the methodology

The description of the subject of this study and the archival sources used, reveal certain restrictions on the methodology and the development of a database. These can be summarized as follows:

1. The sources are limited in number and are very heterogeneous. Due to the dearth of information, we will have to make exhaustive use of what is available. The second problem implies that the prosopographical database we develop must be closely linked to the original archival sources. The database must reveal the idiosyncrasies of the sources and the user should bear in mind the factors determining how the material was handed down. The way in which a source was handed down can be significant: a mortgage deed, for instance, that was inherited until it ended up in an archive, can provide information on the family relations of the successive owners.
2. We cannot decide in advance who belonged to the urban elite and who did not. We mentioned above that *hoofdelingen* without a heritable surname, or whose surname had disappeared by the 17th or 18th century, often escape the notice of historians. This is even more true of the Groningen elite, as councillors with professional names or patronyms are more the exception than the rule in the 16th century. Tallying family names in, for instance, lists of mayores and councillors is common practice. However, it is also a very crude and distorting one, as it takes no account of the kinship in maternal line and relationship by marriage, which were of more importance than family relations in the paternal line. Furthermore, this information must be considered in setting-up a database. Information about individuals who seem insignificant may turn out to be useful to the study. Ideally, therefore, the source materials should be as accessible as possible, time permitting. Even with the limited Groningen sources, however, it is impossible to process the material and make it accessible in such a thorough way as to construct a database of exhaustive biographies of all 15th and 16th century Groningers. That is why we opted for a multi-level database which serves to store raw material as well as to record the processed and logically linked information.
3. It would be a waste of time to conduct a survey on a representative sample because of the limited sources, the small population (in the case of the

hoofdelingen) and the lack of a defined, especially urban, population. For one thing, the non-random nature of the sources passed on is distorting per se. Taking a random sample would enlarge these distortions. In addition, the mutual (family) relationships between the individuals of the population are the very thing that determines the structure of the elite, a structure of relations that is not always visible to the researcher when confining himself to a sample survey. Moreover, the definition of the political urban elite population is not exactly known. It is, therefore, impossible to take a representative sample: before such a sample could be taken from those who 'certainly', 'probably', 'probably not' and 'definitely not' belong to the elite, we have to know the relative proportions and the structure of these four groups, and this can only be determined by studying all candidates. The time that might be gained by taking a random sample would be virtually insignificant, and the danger of distortion would certainly increase.

5. Structure of the database

Before discussing the structure of the files used in this investigation, we should make two observations. First, many of the choices made during the development of the files were historically determined. The approach and focal question of this study evolved during the years in which it was conducted. As a result, the existing files still contain many irregularities from the past. Technical limitations, now a thing of the past, have also left their marks. Moreover, when this study began, personal computers and the accompanying software were hardly available. That is why all data was processed on a mainframe, using programs which we wrote ourselves. After several switches, we are now using a Sun 10/50 workstation running on Unix (SunOs 4.1.3), and connected to other workstations via the Ethernet. The files occupy a total of 0.3 Gigabytes. The programs used, apart from the routines available in the operating system, are the MicroEMACS 3.8m editor and the TeX 3.14 typesetting system. These latter packages can be programmed, thus allowing the user to sort on fields using concordances and thesauruses, and retrieve information from the database. Macro's, routines or 'dirty tricks' have replaced the Pascal programs that we wrote and used during the first few years.

We are not trying to present our system as perfect in this chapter. We merely wish to illustrate that every study has its own requirements with regard to data management -and closely connected to this- the researcher's computer literacy. For this reason, we have kept our summary brief. The structure of the system is illustrated in diagram 1. As can be seen in the

diagram, there are different levels. Once again, we must emphasize that the choices made do not necessarily apply to other types of prosopographical studies.

1. The archival sources (Records and Transcriptions) can be found on the lowest level. Obviously, not all of the texts used have been entered completely. However, it is useful to have easy access to the texts that have been entered, as that simplifies the detailed combination of the different sources. The information from material sources can be distinguished into two categories: context-dependent and context-independent information. The first group includes, for instance, deeds. It is not sufficient to know whether or not someone is listed in a document; the entire content of a deed is needed in order to interpret the occurrence. A tax roll is an example of a context-independent source: if the user wants information on a particular person, he is usually not interested in the entire tax list, but rather in the name and the tax assessment of the individual in question. On this lowest level only the texts of context-dependent sources are directly linked to the higher level. The relevant information from the context-independent sources has been stored in the second level.

2. The first abstraction of the sources takes place on the second level (Briefs). On this level, briefs of context-dependent sources have been stored. They are linked to the complete text and document description on level 1, if available. Briefs of transcribed charters or documents that are otherwise relevant are referred to as well. The structure of these briefs with additional information is uniform for all documents and similar sources. A list of names has been added to each brief, consisting of all entries on the third level that refer to this particular brief.

3. The third level (Indices) consists of three files: indices of context-dependent and context-independent sources and a file of aliases. The first contains the central index of individuals in the index of the previous level. One person may thus occur several times in this file, accompanied by references to different sources in the previous levels and the dates of these sources. To facilitate retrieval and identification, all names are standardized: one form is adopted for interchangeable names, such as Hendrick, Hinric, Hyndryck etc.: in this case, it would become Henrick. A macro in the editor allows the user to check whether the file is still contaminated with deviant spellings, using a thesaurus of these standard names. The name of a person consists of six fields, separated by a slash. For example:

Dr/Cornelis/Gerardi/Begeman/Groningensis/1/

In order to identify someone, the following fields are significant: the second field (first name), the fourth field (surname, patronym etc.) and the

sixth field, a code to distinguish different individuals for whom the second and fourth fields are the same. The code '0' indicates that it has not yet been verified whether the person can be identified with other individuals whose entries correspond with respect to the second and fourth field. Titles, prefixes and suffixes indicating origin of profession (the first, third and fifth field) are not considered for identification. Theoretically, the aforementioned Mr. Begeman would thus (to an interpreting program) be the same as:

/Cornelis/fans/Begeman/van Bremen/1/

but different from:

Dr/Cornelis/Gerardi/Begeman/Groningensis/2/

Identification code 0 is convenient in that not every individual needs to be identified. The file on the third level includes some 250,000 entries, the majority of which will be unimportant to this study. It goes without saying, therefore, that a major selection takes place at this stage, between the actual target group and other individuals. However, this general index is of great importance to the study. Let us take a simple case by way of illustration: say we were searching for information on the hereditary succession of the house of Willem Braeckman. A 1430 document reveals that the respectable, yet insignificant neighbour, Henrick Drost, lives on the north side of the house. In 1450, suppose that it appears that a house in the same street belongs to Anna, the widow of Henrick Drost, whereas Harmen van Bon lives on the south side of her house. It is known that Bon is an heir of the aforementioned Braeckman, the general index provided us with information on the succession of the house in a roundabout way, using the entries of persons who are of no direct significance to the elite studied. It is information such as this which is so important in interpreting sources, that makes a general index vital.

The second file contains information from the context-independent sources. Its structure depends on the source. For instance:

\bsc /Deric/de/Vries/1/Ebbingestrate/43/5-0-0/

means that: Deric de Vries (with identification code 1) was taxed 5 Emders guilders in the *burgerschaiting* (tax roll of Groningen burghers) of 1535, 2) lived in Ebbingestrate and 3) is mentioned as no. 43 on the list. This can be significant if the sequence of the list is relevant, as it is in this case because the entries of the roll are arranged in the order of the actual sequence of the houses. However, this information is not needed often enough to have the complete list available on-line. The string:

\wsz /Ingele/Clapwijck/0/ARs/50/

means that Ingele Clapwijck, who has not yet been identified with namesakes, used, according to the Winsumer and Schaphalster Zijlvest's taxation

roll in 1553, 50 yokes of land in Roodesch. The place name Roodeschol has been coded as VRs so as to keep the volume of the database within manageable limits, which would have been impossible if names of parishes, functions and other frequent words had been entered in full. The databases have been structured in such a way that when the queries are printed, the TeX program replaces the abbreviations with the full name.

In addition, this level contains an alias file, in this file persons who occur under different names are linked to the names in the central index. Examples are women who occur under their maiden name, as well as under the names of their respective husbands and individuals who occur under their occupational titles, patronyms, toponyms or under different surnames. Obviously, the individuals listed in this index have an identification code higher than 0.

For practical reasons, the information is usually divided into several files. Some of those can be merged into one file when this is required to make certain queries.

4. The highest level (level 4 in diagram 1) contains the actual database with information on the most relevant individuals. The names are divided into 6 fields, as on the previous level. However, on this level, every individual occurs only once. Moreover, the names with an identification code '0' have not been included. The following information has been recorded here for every individual: possession of real estate, family relations, positions and genealogical data, education etc. Given the incomplete nature of the archival sources, it is not surprising that most fields are empty. Consequently, making queries on this level is possible only to a limited extent and less worth the effort due to the disparate archival sources. Most questions of a more or less statistical nature are, therefore, applied to the context-dependent files in the previous level. Thus, results of comparing different sources are not only to be found in the highest level of the database. While on this level the biographical data of persons are stored, the information on lower levels is used in order to examine their relationship to, for instance, property in a larger context.

6. Conclusions

The system described above has a number of indisputable advantages:

1. It has been adapted to the situation of the local archival sources, the methodology and the individual needs of the researcher.
2. The multi-level design takes into account the limited amount of time set for this study and the unpredictability of the time required to interpret and

process the archival sources. Depending on these factors, the researcher may restrict the study at a later stage to a smaller population or to a limited number of questions.

3. It guarantees the researcher all of the access to the sources that he needs in order to answer certain questions : sometimes interpretation requires consulting a brief or a transcription.
4. The problems of identification are mostly dealt with.
5. Because the structure of the entries was designed, among other things, to allow use of the typesetting system TeX, control of syntax and word processing for presentation of interim and final results takes no extra effort. In the past this has shown to be a very useful feature.
6. The information was stored in a structured way, thus enabling linkage with other databases, if desired. This also allows the researcher to take a more flexible approach. Because the editor can be programmed and it recognizes the used syntax, the structure could easily be altered over the last few years in order to meet new requirements. In most cases restructuring did not cost much time.
7. The size of the files is unlimited.

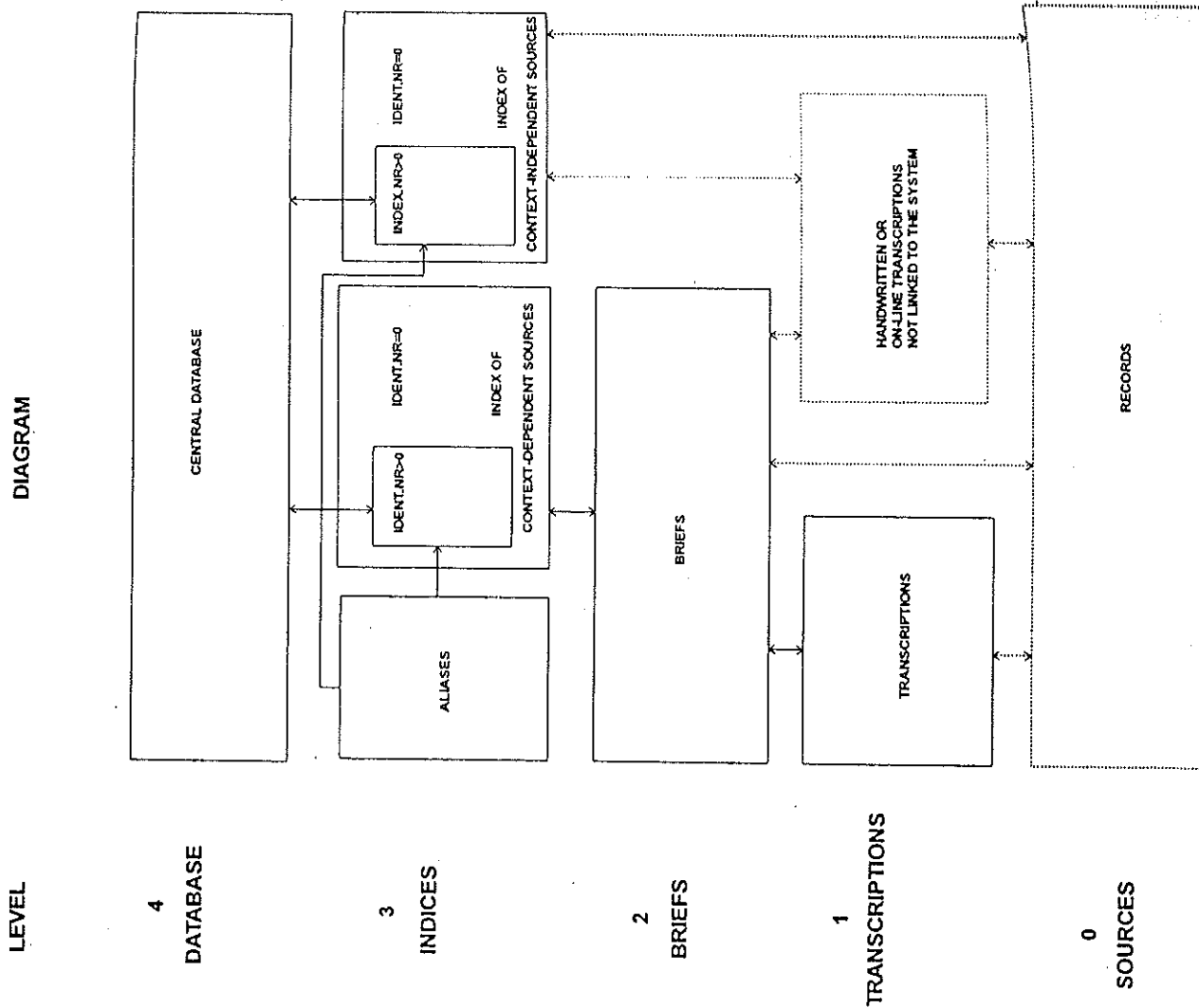
There are also a few disadvantages to the system:

1. Because the system was created for one particular study and one particular researcher, it is not necessarily suitable for other studies or for use by other researchers.
2. The size is considerable, which puts constraints on memory space and speed of the computer.

There is no ideal system for all forms of prosopographical research. Prior to each study, the archival sources, the questions asked, the idiosyncrasies of the population, the availability of computers and the (programming) experience of the researcher must be carefully considered in choosing an already existing software package, or writing one's own programs, if possible. Obvious as they may be, this preparation is often left undone. If prosopography is to shed its image of proving the obvious, flexibility in this respect is vital.

Diagram :

Arrows indicate links by means of name and identification code and/of denotation of an archival source.
 References to data outside the system itself are represented by dotted lines



Prosopography and source oriented data processing with Kleio

GODFRIED CROENEN

1. Introduction

For several years I have been engaged in doctoral research into the acquisition of power and family structures in the Brabant nobility for the period from around 1100 to about 1350. In this context I have concentrated on the Berthout family. In addition, those families related by marriage and the families of vassals have been the subject of study as well. The aim of my research has been to gain an insight into the structure and the social functioning of a noble family in the period studied and into the ideological aspects of it as well as the way in which such a family functioned as a power base.

My research draws to a great extent on information about persons. In the first place, of course, it is important to draw up good genealogies. For this purpose, as much information as possible is gathered about family ties, inheritances, transactions between various persons and so on. In addition, information is collected regarding the holding of seigneurial titles, official posts, the use of seals and types of seals. Information is also gathered on functions such as acting as a guarantor, endorser or witness for other nobles or lords, sitting on the ducal council or being part of the ducal feudal court. In the case of the Berthout family, there is data on over a hundred persons. Data is also collected about families related by marriage, other noble families in Brabant and about families and persons connected with the Berthouts through personal ties or business dealings. The source material consists mainly of deeds. These often appear to contain little relevant information. Nevertheless, on close analysis, the records are shown to offer a great deal of scope because they are so numerous, containing about 1300 texts alone, referring directly to the Berthout family. The method followed is not strictly prosopographical, but depends to a large extent on prosopographical data.¹

In Section 2 of this text, I shall first try to give a more precise description of my requirements regarding the use of computer technology and indicate a number of problems which immediately present themselves in this respect. Section 3 will indicate how I approach the work in practice and the

¹ Cf. MILLET (1985).