Karina van Dalen-Oskam
(The Hague)

MAPPING THE ONYMIC LANDSCAPE

Abstract. Several researchers have argued for an analysis of names in literary texts in the context of the other names in the same texts, which implies the analysis of *all* names in a text, oeuvre, genre or time period to establish their exact function(s) in that text, oeuvre, etc. Mapping the onymic landscape is time consuming but information technology could help to make the analysis of all names in a designated corpus a lot easier. The paper introduces Autonom, a digital workbench for (collaborative) literary onomastic research and shows the tool's possibilities and potential by describing the first steps in the analysis of the names of the important Dutch novel *Nooit meer slapen* (1966) by Willem Frederik Hermans (1921-1995).

1. *Introduction*

The discipline of literary onomastics has undergone significant developments in recent decades. Several researchers have argued in important publications for an analysis of names in literary texts in the context of the other names in the same texts, which implies the analysis of *all* names in a text, oeuvre, genre or time period to establish their exact function(s) in that text, oeuvre, etc.

There is, however, a practical problem in this respect: mapping the onymic landscape is terribly time consuming if done in the traditional, manual way and this probably keeps researchers from doing it. But the gigantic developments in information technology and their impact on the linguistic and literary research communities during recent decades could help to make the analysis of all names in a designated corpus a lot easier. In this paper, I introduce Autonom – a digital workbench for literary onomastic research – and show the tool's possibilities and potential by describing the first steps in the analysis of the names of what in the Netherlands is one of the most influential novels of twentieth-century Dutch literature, namely *Nooit meer slapen* (1966) by Willem Frederik Hermans (1921-1995).

Nooit meer slapen ("Never sleep again") is written from the point of view of Alfred Issendorf, a young geologist who has just started work on his thesis on the landscape of Finnmark and is looking for evidence of

meteorites. He is anxious to prove himself personally, socially and scientifically as being the equal of his father, a botanist who died young during a scientific expedition in the Swiss mountains. Alfred joins an expedition to Finnmark with three geologists from Norway. He befriends one of them (Arne) but feels threatened and tricked by the other two (Mikkelsen and Qvigstad). Alfred proves not to be up to the harsh conditions of the landscape and fails to prevent Arne from falling to his death. He does not acquire any scientifically interesting information for his thesis and returns home a disappointed man.

This novel became a cult book for many of the young Dutch men and women who grew up in the sixties and seventies. W. F. Hermans is now (still, or perhaps increasingly) considered as one of the most important – if not *the* most important – Dutch novelist of the twentieth century. Recently (4 November 2005), the first volume of a scholarly edition of his collected works was presented to Dutch Crown Prince Willem-Alexander during an impressive laudatory gathering in the Nieuwe Kerk in Amsterdam. The 24-volume series is edited by the Huygens Instituut of the Royal Netherlands Academy of Arts and Sciences.

Nooit meer slapen was translated into Swedish in 1968 (*Aldrig mera sova*, transl. Brita Dahlman), German in 1982 (*Nie mehr schlafen*, transl. Rosemarie Still, 2nd. ed. 1986) and Norwegian in 1992 (*Aldri sove mer*, transl. Eva Paasche). In 2002 a new German translation was published (*Nie mehr schlafen*, transl. Waltraud Huesmert; paperback ed. 2004), and two years later the Estonian translation appeared (*Igavene uni*, transl. Kerti Tergem). It is unknown when an English translation will be published, but it is certain that more and more of Hermans's works will become available in translation during the coming years.

Many readers feel that the names Hermans uses for his characters are very idiosyncratic and recognizable. The aim of my research is to find out why Dutch readers think that they can denote a literary name as "typically Hermans". The onomastic analysis of *Nooit meer slapen* is the first step towards answering this question. In this paper only the very first part of this research is described. The focus is on the selection of the data, the methodology to be developed and the kind of new research possibilities computer assistance could provide. A thorough analysis of the names will have to wait for another occasion.

2. *Autonom*

Autonom is a web application under development that provides a

framework for the computer-assisted analysis of texts (http://autonom. huygensinstituut.knaw.nl).The first tool to be developed within this framework was a name parser for Dutch. This parser "reads" a digital text provided by the researcher and returns a list comprising all the words that mightbe names. In modern Dutch, the initial capitalization of a word occurs only when a word is the first word of a sentence, when it is a name or when itis part of an acronym. The rule used for the parser is based on the assumption that a word with an initial capital is most probably a name if it does not occur in the same text without the first character being capitalized. Only very occasionally is a name missed by the parser, for instance for the male first name *Ben*, which would be ruled out by the first-person singular indicative present *ben* of the verb *zijn* ("to be"). There are a substantial number of words returned that happen to occur only once and then at the beginning of a sentence, but these can be very easily recognized in the list.

This list is then edited by the researcher, who can reject or confirm an item as being a name. The accepted names can be tagged with appropriate labels for further analysis. I will describe this procedure for *Nooit meer slapen*. A digital file of the novel was created using a scanner and an OCR (optical character recognition) technique. This text file (in ASCII) was uploaded into Autonom. The text was not corrected, because OCR mistakes are irrelevant after all names have been correctly labelled (e.g. lemmatized). For copyright reasons, the text can be viewed only by the researcher who uploaded it. Then the name parser was activated, resulting in an alphabetical list of all possible names. Figure 1 shows part of this list, with the item *Amsterdam* activated showing the tags added to it in the first phase of the analysis. Figure 2 shows how one can take a look at all occurrences of a word form as key words in context, here for the name *Amundsen.*
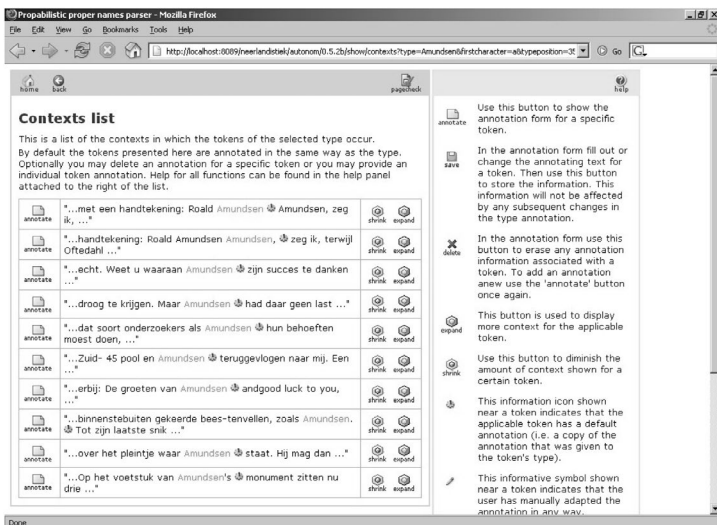
In Figure 3, the context of the first occurrence of *Amundsen* is further expanded and the labels added by the researcher are shown. *Nooit meer slapen* has a lot of references to celebrated mountaineers and artic/antarctic explorers and this is reflected in the chosen labels. This is shown in Figure 4, where a label registers the fact that *Scott* and *Shackleton* are mentioned in combination in the text. Because a researcher will hardly ever know in advance what to expect in a text and in the process of analysis, Autonom does not prescribe which labels to use. The researcher is free to use his or her own labels.

In the next phase of the technical development of Autonom, a filter will enable the researcher to view all given labels and allow the possibility to change or elaborate on them. The aim of this, of course, is to create the

possibility for an easy and efficient quantitative and qualitative compari-
son of the names. Although the web application does not yet contain the
tools for this sort of visualization of the added labels, a first start with this
type of research will be made in the following section.
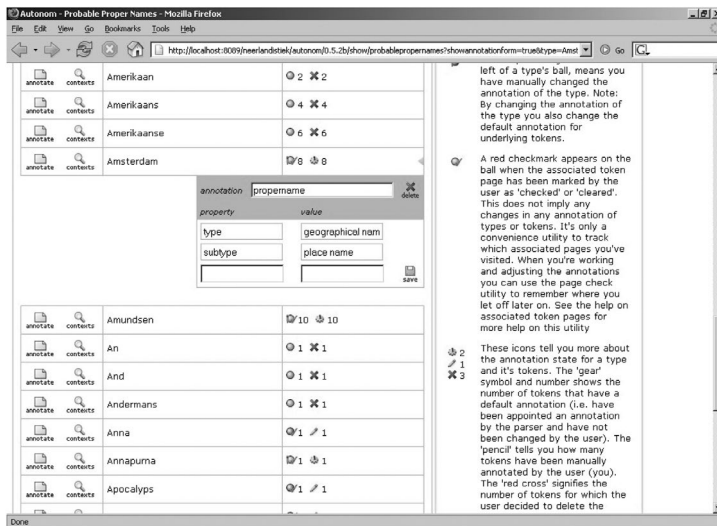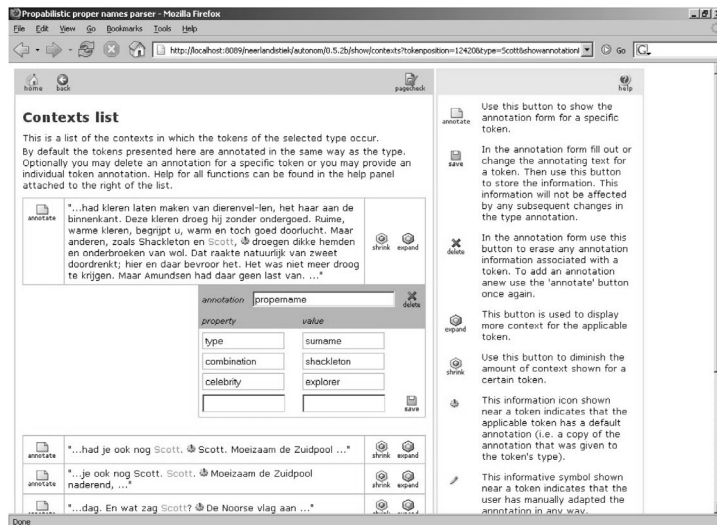

[Fig. 1: Alphabetical list of possible names returned by the name parser]

[Fig. 2: Context list for the name *Amundsen*]
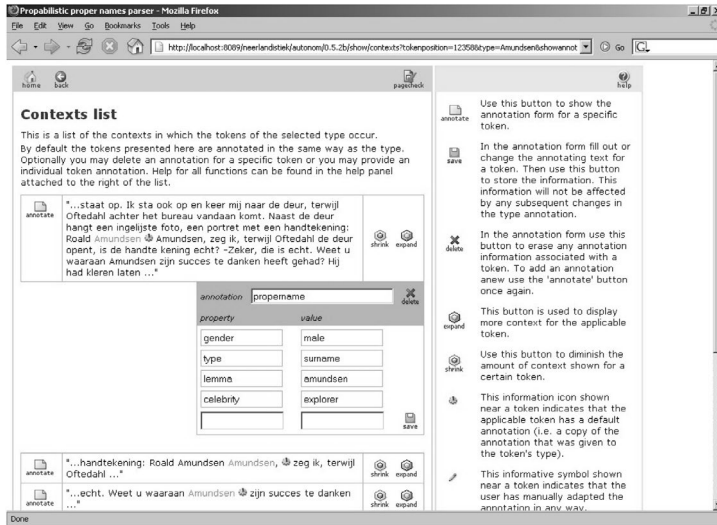
[Fig. 3: The expanded context for the first occurrence
of the name *Amundsen* and the tags added by the researcher]

[Fig. 4: The expanded context for the first occurrence
of the name *Scott* and the tags added by the researcher]

3. *Mapping the onymic landscape of* Nooit meer slapen: *a peek at the future*

If one wants to know what is special about a name or the names in a literary work, oeuvre, genre, etc., one first has to establish what is normal. This implies that a lot of work is still to be done and that at the moment it is impossible to reach definitive conclusions about the names in *Nooit meer slapen*. But finding out what is normal can only be done with the elaborate assistance of information technology in general and, hopefully, with Autonom in particular. What could we learn by labelling all names in a text with help of Autonom?

Total number of names (incremental)
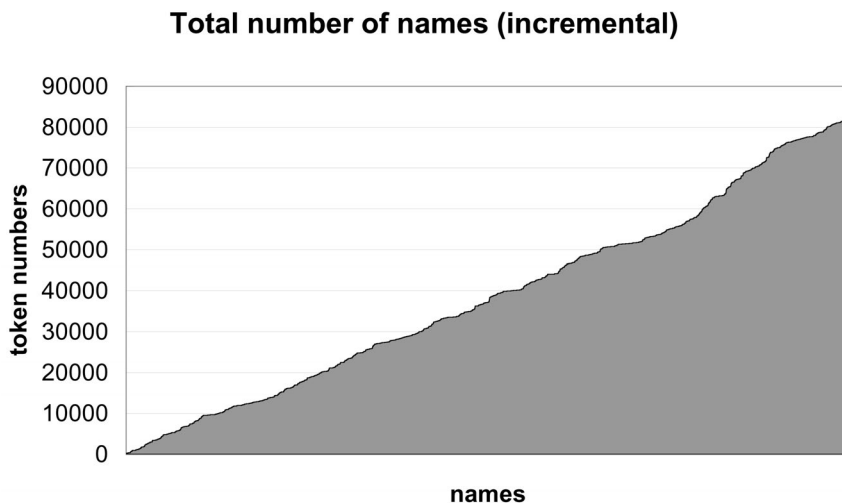
[Fig. 5: The occurrence of names (incremental) throughout the text of *Nooit meer slapen*. The text consist of about 82,000 words (tokens).]

We could find out what kind of spread names have in the text and compare this with as many other texts as possible. Figure 5 shows a simple graph visualizing the occurrence of names throughout *Nooit meer slapen*. This graph cannot yet be made with help of Autonom. Excel was used to experiment with, but the program proved very laborious and impenetrable (although Excel has an enormous number of possibilities, it has a steep learning curve).

In this experimental graph, one can see that names occur in a relatively

steady flow throughout the text. Although we have not yet been able to test this on a large text corpus, it is to be expected that this is the normal situation for a literary text. If part of the graph were flat or distinctly steeper, this could point us to parts of the text where something exceptional is going on and which thus deserve closer investigation.

The next step could be a graph in which reoccurrences of names are neglected, thereby only showing the introduction of "new" names in the text. Is the intuition true that most new names are introduced in the first part of a text? What is happening in a text that deviates from this structure? Related to this is the number of different names in a text. It seems logical that the number of different names or even the choice of name types (e.g. only first names) could highlight the intimacy of a story or could be one of the means by which an author emphasizes the smallness of the "world" he or she is describing. What we hope Autonom will also be able to do in the future is show us the area in a text in which certain names or name types occur, thus visualizing for instance changes in focalization and/or the importance of characters in the text.

With Excel it was relatively easy to make a pie chart showing the three name types that occur most frequently in *Nooit meer slapen* (as tokens, not as types). In Figure 6, the occurrences of names not of these types are grouped together as "other names" (in the smallest part).
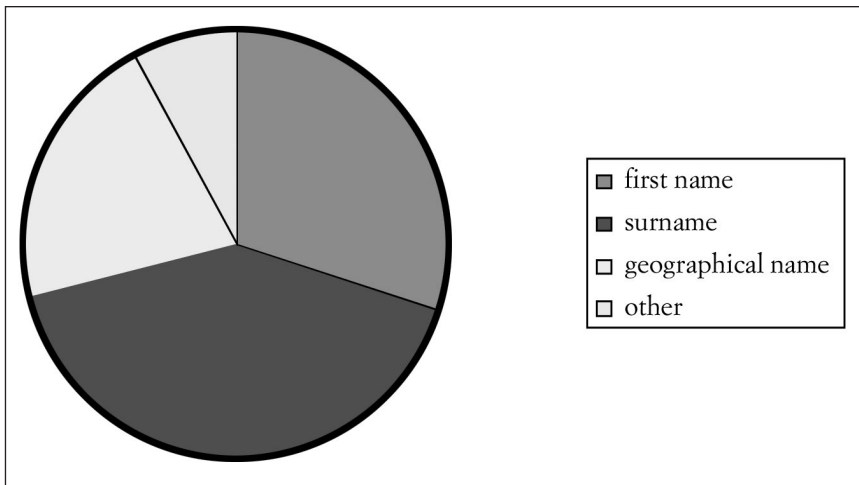
[Fig. 6: The three name types that occur most frequently in *Nooit meer slapen*]

The total number of occurrences per name type is:

First name	497
Surname	678
Geographical name	343
Other	134
Total	1652

These numbers could of course differ per text, author, genre, etc. - but by how much? We do not know yet. Could we link the relative use of female and male first names to a certain text genre, time period or culture? In *Nooit meer slapen*, female first names are represented by 56 tokens, denoting 15 different persons; 29 tokens denote *Eva*, sister of Alfred. Male first names occur 435 times, denoting 39 different persons. The token *Arne* occurs 326 times. Hermaphrodite *Ymir* is mentioned four times. This highlights the relative absence of the feminine in Alfred's world.

Focusing on the surnames, we see that *Nooit meer slapen* contains 60 different ones. Together they occur 678 times. The most frequent surnames are:

Nummedal	143
Qvigstad	143
Mikkelsen	129
Sibbelee	66
Hvalbiff	38
Oftedahl	35
Brandel	14
Amundsen	11
Scott	8
Grieg	7
Livingstone	6
Flintstone	4
Jordal	4

A lot can be learned from this list. Alfred's friendship with Arne Jordal is reflected by the predominant usage of his first name and the scarceness of his surname. Alfred's hate of the other two Norwegian members of the expedition is evinced in his use of only their surnames, and mostly as an extra-threatening *pair* of surnames: "Mikkelsen and Qvigstad". At the top of the list is Professor Nummedal; again, this a person Alfred has a lot of problems with. He feels that the Norwegian professor has deceived him

and has given Mikkelsen and Qvigstad an unfair advantage by giving them aerial photographs that he had told Alfred did not exist. The name Nummedal is the best example of a "typical Hermans" name. This surname does not exist in Dutch, but sounds rather a lot like *niemendal*, which means "worthless, a trifle" - which sums up Alfred's opinion of the professor quite nicely. But there is more to the name: scholars of Norwegian names have informed me that Nummedal is not only a surname in Norway, but also the name of a well-known Norwegian archaeologist (many thanks to Botolv Helleland and Benedicta Windt for this information).

As mentioned, several famous explorers are repeatedly mentioned in the novel when Alfred muses about their feats, although the list above shows that Amundsen, Scott, Livingstone and certain others are mentioned only a few times (Flintstone is another story, but the name itself is of course nicely chosen for a novel about a geological expedition). The names of these explorers have a clear function in describing the world as seen from Alfred's perspective: persons and places that have a special meaning for Alfred are foregrounded by being mentioned by one of the characters or by the narrator. In the novel, these names thus highlight Alfred's passionate ideas about exploration: he regards explorers like these as heroes. Although he would like to think that his father was and that he himself will be part of that select group, he finds out that they both are failures in this respect.

All names related to exploration have been labelled with the help of Autonom. The next step will be to investigate more closely the way in which in *Nooit meer slapen* Hermans uses these names to foreground the motifs and themes (or how the names *are* motifs or visualized themes). Could Alfred's mental development be read in a certain change in the use of names associated with exploring or other themes? Is a change in this respect related to other changes in the use of names throughout the text? These sorts of questions could hardly be asked before, but could provide a deeper insight into the ways names function or are experienced in a literary text.

4. *Conclusion*

At the moment, Autonom can be used for only a few basic kinds of analysis. However, the application has the potential for a lot more, as the above has shown. Apart from the already-mentioned desire for a tool that will generate graphs of all kinds of added information, it would be useful to be able to share analyses of texts or text corpora with other researchers.

If different researchers were to analyse different texts and provide their colleagues with information about the method they used to label their texts, they could compare results and, through discussion, substantially further literary onomastics and the development of new methods and tools.

Autonom could become a virtual workplace for literary onomasticians from all over the world. In addition to joining forces scientifically, however, they would also have to join forces financially. New tools have to be developed and added (first and foremost, name parsers for languages other than Dutch), and a lot of work still has to be done in order for visualizations and quantifications to be performed easily. Any researcher who is interested in cooperating with the Huygens Instituut in this respect is invited to discuss the possibilities for collaboration and joint funding applications.

Note

Bibliography

F. Debus. *Namen in literarischen Werken: (Er-)Findung - Form - Funktion.* Mainz, 2002.

K. van Dalen-Oskam. *Vergleichende literarische Onomastik.* A. Brendler und S. Brendler (Hrsg.). *Namenforschung morgen: Ideen, Perspektiven, Visionen.* Hamburg: Baar, 2005, p. 183-191. (*Comparative Literary Onomastics*; English translation available at www.huygensinstituut.knaw.nl through "Medewerkers" > "Karina van Dalen-Oskam" > "Publicatielijst".)

K. van Dalen-Oskam. *Dutch literary onomastics. Past, present and future.* Eva Brylla, Mats Wahlberg, in collaboration with Vibeke Dahlberg, W.F.H. Nicolaisen (Eds.). *Proceedings of the 21st International Congress of Onomastic Sciences, Uppsala, 19-24 August 2002.* Vol. 1, Uppsala, 2005, p. 398-406.

K. van Dalen-Oskam & J. van Zundert. *Modelling Features of Characters. Some Digital Ways to Look at Names in Literary Texts.* "Literary and Linguistic Computing" XIX (2004), 3, pp. 289-301.

K. van Dalen & J. van Zundert. *Finding Names with the Autonom Parser.* "DigiCULT.info" 6 (December 2003), pp. 47-49. (http://www.digicult.info/downloads/dc_info_issue6_december_20031.pdf).

W. NICOLAISEN. *Names in English Literature*. E. Eichler a.o. (eds), *Name Studies. International Handbook of Onomastics,* Vol. I, 560-568 (Handbücher zur Sprach- und Kommunikationswissenschaft Band 11.1) Berlin/New York, 1995.

M. PALACAS. *The Importance of Being "Ernest" in Mary Shelley's Frankenstein: A Study in Literary Onomastics*. "Names: a Journal of Onomastics" LIII (2005), 3, pp. 203-234.

I. SOBANSKI. *Die Eigennamen in den Detektivgeschichten Gilbert Keith Chestertons. Ein Beitrag zur Theorie und Praxis der literarischen Onomastik.* (Europäische Hochschulschriften Reihe XXI (Linguistik) Bd. 218; Diss. Leipzig, 2000), Frankfurt am Main etc., 2000.