Quo Vadis Lexicography at the Institute for Dutch Lexicology?

Jeannine Beeken

Instituut voor Nederlandse Lexicologie, Leiden/Antwerp

In this paper, we will first introduce the Instituut voor Nederlandse Lexicologie (INL, Institute for Dutch Lexicology). We will present an overview of the INL-dictionaries online, being the Oudnederlands Woordenboek *(ONW, Dictionary of Old Dutch; ca. 475 – 1200), the* Vroegmiddelnederlands Woordenboek *(VMNW, Dictionary of Early Middle Dutch; 1200 – 1300), the* Middelnederlandsch Woordenboek *(MNW, Dictionary of Middle Dutch; ca. 1250 – 1550), the* Woordenboek der Nederlandsche Taal *(WNT, Dictionary of the Dutch Language; ca. 1500 – 1976), the* Etymologisch Woordenboek van het Nederlands *(EWN, Etymological Dictionary of Dutch) and the* Algemeen Nederlands Woordenboek *(ANW, Dictionary of Contemporary Dutch; 1970 - 2018). Thirdly, we will present the Taalbank Nederlands (Language Bank Dutch) and its main tasks. Finally, we will elaborate on three U-turns, namely a first U-turn: from manual labour and printed material to computational linguistics and the internet, a second U-turn: from single functionality to multiple functionality, a third U-turn: from stand-alone product to spin-offs, linking and integration. We will finish with some thoughts and ideas answering the following question: quo vadis lexicography at the Institute for Dutch Lexicology?*

1. The Institute

The Instituut voor Nederlandse Lexicologie (INL, Institute for Dutch Lexicology) was instigated on September 14th 1967; nearly 43 years ago. It is located in Leiden and has had an association with Antwerp since 2004. The Institute is Dutch-Flemish and receives its funding from the governments of the Netherlands and Flanders. The first assignment in the late 60s was to end, after more then 100 years of lexicographical work, since the 1860s, the *Woordenboek der Nederlandsche Taal,* i.e. the Dictionary of the Dutch Language (WNT). The dictionary was eventually finished in 1998, after nearly 150 years, covering 5 generations of lexicographers; it contains 39 volumes and 1 supplement. In 2001 another 3 volumes with supplementary information were published. The WNT ranges over a period from 1500 till 1976, i.e. from the 16th till the 20th Century. It contains 400.000 lexical entries and 1.7 million quotations. In January 2007 the WNT came free available online; it was the first of the INL-dictionaries online and it has to this point more than 70.000 individual subscribers. That is without the subscriptions from libraries and universities.

At the moment, 50 employees are working at the Institute. More than half of them are academics, the other personnel are supporting staff, e.g. assistants and technical staff such as system analysts and programmers. Half of the scientific staff has a PhD in either linguistics or computational linguistics. The overall budget of the Institute is 2.8 million euro a year; the TST-Centrale (HLT Agency, Agency for Human Language Technology), which is the Dutch-Flemish central point for maintaining, supporting and distributing all the language and speech technology materials that is funded by governmental institutions has a yearly budget of about half a million euro. The Institute is an independent foundation which since 2008 falls under the overall responsibility of the Nederlandse Taalunie (Dutch Language Union). The Institute likes to profile itself as the Treasurer of the Dutch Language, at the moment, as far as the vocabulary of the language is concerned, of course.

2. Dictionaries

2.1. *Oudnederlands Woordenboek* (ONW)

Which are the treasures of the INL? I'd like to give an overview of all of our dictionaries. I'll start with the *Oudnederlands Woordenboek* (ONW, Dictionary of Old Dutch), ranging over a period from ca. 475 till 1200. The corpus of Old Dutch, on which the dictionary is based, contains about 50.000 word forms or tokens in different spellings, half of which are toponyms. The dictionary itself contains about 4500 entries. The examples and quotations are taken from the corpus of Old Dutch; every quotation has been translated into modern Dutch, trying to save the original syntax as much as possible.

The work on the dictionary started in 1998. 10 years later, in 2008: the dictionary of Old Dutch was a fact. It is the first dictionary constructed at the Institute which will not be available in printed form. The dictionary came available online in May of this year (http://onw.inl.nl); the corpus is available since 2006. As said, the *Oudnederlands Woordenboek* is only available online, both as a stand-alone product and as part of the integrated application *Historische Woordenboeken van het Nederlands online* (Historical Dictionaries of Dutch online, http://gtb.inl.nl), on which I will elaborate later on. Very important, however, are two types of standardisation that have been put through. First of all, all of the Old Dutch entries are classed under one selected spelling form, linked of course to the original spellings. In that way all the lexical-semantic information could be put under one head form. Secondly, each Old Dutch entry has been translated into a modern equivalent. This was necessary in order to enable the users, who are not always experts, to look up the information, even though the original spelling is unknown to them. Keeping the original spellings also, however, allows the user to directly click on a word in a text and be directed to the corresponding entry in the dictionary. Moreover, since the *Oudnederlands Woordenboek* is part of the INL-dictionaries online, by using modern Dutch equivalents as an umbrella, the users can look up information about a word spread over a time span from 500 AD till modern times. Besides that, for each entry, the relevant reference information is given in order to establish a link with the other historical dictionaries online. In case of a reconstructed form that is not attested in the original material, an asterisk is placed to notify the users of its peculiar status.

Unless new Old Dutch material is discovered, the material used for the corpus as well as for the dictionary can be seen as fully complete. In order to make the information internationally more accessible, German, English and French equivalents of the head entries will possibly be added in a later stage, once funding is found.

2.2. *Vroegmiddelnederlands Woordenboek* (VMNW)

The *Vroegmiddelnederlands Woordenboek* (VMNW, Dictionary of Early Middle Dutch) describes the Early Middle Dutch of the 13th Century. It contains ca. 26.000 entries amongst which are toponyms and anthroponyms. The work on the dictionary started in 1988 and it was finished in 1999 and published in 2000. The dictionary contains 4 volumes but is in printed form no longer available on the market.

The dictionary is completely based on a corpus of texts written in the Middle Dutch of the 13th Century, the so-called Corpus Gysseling. That corpus was published from 1977 untill 1988. Since May 2007, the *Vroegmiddelnederlands Woordenboek* has been available online as a stand-alone product (http://vmnw.inl.nl) and as part of the application *Historische Woordenboeken van het Nederlands online* (Historical Dictionaries of Dutch online,

http://gtb.inl.nl). All of the entries are linked to their modern Dutch counterparts; the corpus itself has also been digitised.

2.3. *Middelnederlandsch Woordenboek* (MNW)

The *Middelnederlandsch Woordenboek* (MNW, Dictionary of Middle Dutch) describes the language of the 13th till the 16th Century. It contains 11 volumes and ca. 75.000 entries; the first volume was published in 1885; the last one in 1929. It took the lexicographers about 45 years to finish it. In 1941 two additional volumes were published. The dictionary is no longer available in printed form. The electronic version of the dictionary has been available online since mid December 2009 (http://mnw.inl.nl). It completed the series of the *Historische Woordenboeken van het Nederlands online* (Historical Dictionaries of Dutch online, http://gtb.inl.nl).

2.4. *Woordenboek der Nederlandsche Taal* (WNT)

I already have given some information about the *Woordenboek der Nederlandsche Taal* (WNT, Dictionary of the Dutch Language) which took about 150 years to finish. It describes the Dutch language from the 16th till the 20th Century. It contains ca. 400.000 entries and it was the first of the historical dictionaries online (http://wnt.inl.nl); it has been available on the internet since January 2007.

The *Woordenboek der Nederlandsche Taal* has links to the *Etymologisch Woordenboek van het Nederlands* (EWN, Etymological Dictionary of Dutch), about which I will talk shortly later on, the *Etymologiewoordeboek van Afrikaans* (EWA, Etymological Dictionary of Afrikaans) and some botanical dictionaries. On the short term, the entries will also be linked to the dictionaries of the Brabantic, Flemish and Limburg dialects. In the summer of 2010 a link is established between the *Woordenboek der Nederlandsche Taal* and its Frisian counterpart, the *Woordenboek der Friese Taal* (WFT, Dictionary of the Frisian Language). The same could be achieved with the Afrikaans counterpart (the WAT). Worth mentioning is that every entry of the WNT is translated into its modern Dutch variant.

2.5. *Etymologisch Woordenboek van het Nederlands* (EWN)

The *Etymologisch Woordenboek van het Nederlands* (EWN, Etymological Dictionary of the Dutch Language) has been completed in November 2009. In November, the fourth and last volume of the dictionary was published in printed form; the first volume dates from 2003, the second from 2005 and the third from 2007. It contains the etymology of about 15.000 contemporary words. The etymological dictionary has also been incorporated in the online application *Historische Woordenboeken van het Nederlands online* (Historical Dictionaries of Dutch online, http://gtb.inl.nl), where it is integrated in the *Woordenboek der Nederlandsche Taal*.

2.6. *Algemeen Nederlands Woordenboek* (ANW)

The *Algemeen Nederlands Woordenboek* (ANW, Dictionary of Comtemporaruy Dutch) describes the contemporary Dutch language as used in the Netherlands and Flanders. It ranges over a period from 1970 till 2015. The project started in 1998 and will be finished by the end of 2018; it runs over a period of 20 years. In the preparatory phase much attention was paid to the following two tasks: the creation of a text corpus of ca. 100 million words, suitable for the purposes, and the development of the concept. Part of the concept was the identity of the dictionary, the identification of the lexical classes, the development of a lexicographical work station or dictionary writing system and the identification of the microstructure of the entries.

First of all, from the very beginning the *Algemeen Nederlands Woordenboek* has never been set up as a dictionary meant for printing. It was seen as a corpus-based explanatory dictionary of modern Dutch that will only be available online. It was never set up to be an alphabetically organised dictionary of which volumes would be published throughout the period of 20 years. Therefore, the classification of the material is not alphabetic, but according to the morphosyntactic-semantic classifications and domains.

The dictionary will contain about 70.000 entries and 250.000 subentries, organised as word families, i.e. a headword together with its derivations and compounds. The semantic classes are characterised by frames, i.e. a set of slots and fillers. The dictionary will not only contain phonological and morphosyntactic information, but also definitions, of course, and semagrams. The latter are completed frames representing the knowledge, also the encyclopaedic knowledge that is connected with a specific word.

From the start, attention was paid to the back-end, which means the functional design of the application and the front-end which stands for the graphical design of the web interface for the user. Besides that, a dictionary writing system has been developed for the lexicographer who completes the information for the different information types which are linked as and saved in a relational database.

The first results of the ANW have been made available online in December 2009 (http://anw.inl.nl). It is possible for the users to explore the dictionary in different ways. The two most important ones are the semasiological and the onomasiological search. The semasiological profile of the lexical items is the set of meanings together with their internal structure and their mutual relations. The user, when using the semasiological search types in a word or expression, will be guided to the set of meanings etc. This is the search that is standard for most dictionaries online.

The onomasiological search is based on one or more characteristics, corresponding to one or more fillers of the semagrams. It will lead the user to the correct word. For example, typing in 'animal, gives milk' will lead to *koe* 'cow', *geit* 'goat' and *schaap* 'sheep'. Typing in 'animal, gives milk, plump' will only lead to *koe* 'cow'. The more fine-tuned the question, the more limited the set of answers. This is a fairly innovative way of exploring online dictionaries. For every day users, this way of exploring a dictionary might seem quite familiar; indeed, they are used to filling out crosswords. Moreover, through the future links to be established in the overall database, on which I will elaborate next, the users will automatically be guided to the corresponding entries in all of the historical INL-dictionaries immediately.

3. Overview

By the end of 2009 the following dictionaries were completed and were made available online: the *Oudnederlands Woordenboek* (5th till 12th Century, http://onw.inl.nl), the *Vroegmiddelnederlands Woordenboek* (13th Century, http://vmnw.inl.nl), the *Middelnederlandsch Woordenboek* (13th till 16th Century, http://mnw.inl.nl), the *Woordenboek der Nederlandsche Taal* (16th till 20th Century, http://wnt.inl.nl) and the Etymologisch Woordenboek van het Nederlands. It will have taken approximately 4 years to put all of these dictionaries online. That means that the Dutch Language from its earliest contestations in the 5th Century till modern times will be described and will be made available on the internet, for free for all users without commercial interests.

As said, also the first results of the contemporary *Algemeen Nederlands Woordenboek* have been put online (http://anw.inl.nl). So, before the end of 2010, when the users type in the word for 'cow', which in modern Dutch is *koe*, they will see all of the information linked to *koe* in the ANW and the WNT, to *coe* in the MNW and the VMNW and to *kuo* in the ONW.

All of the historical dictionaries are part of the project *Historische Woordenboeken van het Nederlands online* (Historical Dictionaries of Dutch online, http://gtb.inl.nl), developed by the Taalbank Nederlands, i.e. the Language Database. Once the *Algemeen Nederlands Woordenboek* has completed the series, the INL will present the results as the series *Woordenboeken van het Nederlands online* (Dictionaries of Dutch online). Next, I will introduce the Taalbank Nederlands, the Language Database of Dutch.

4. Taalbank Nederlands

Once the descriptive work is completed, the dictionaries are handed over to the Taalbank Nederlands (Language Database of Dutch) for further maintenance, integration and disclosure. The Taalbank Nederlands has three major components: a dictionary component, a corpus component and a lexicon component.

The first one I elaborated on previously. The second one, the corpus component, contains all of the source material, i.e. the texts, on which the dictionaries are based. Besides collecting the material, the Taalbank Nederlands also annotates the texts by means of lemmatising and categorising the words or tokens. There is a corpus of Old Dutch texts from the 5th till 12th Century, a corpus of 13th Century Dutch, a series of Bibles ranging from the 15th till 17th Century, modern texts from the late 20th Century and the early 21st Century.

As can be noticed, there are no digitised texts available at the Taalbank Nederlands from the 14th till 15th Century and from the 17th till 20th Century. These gaps are, for the time being at least, filled by the many quotations that can be found in the *Middelnederlandsch Woordenboek* and in the *Woordenboek der Nederlandsche Taal* (about 1.7 millon). How can both these gaps in time be filled?

Firstly, the Taalbank Nederlands tries to acquire the necessary material from third parties. It has for example set up projects with the Royal Library in The Hague to make use of their huge amount of digitised historical material, ranging from the 16th Century onwards, 25 billion words in total. From two research institutes, 14th Century material will be made available during 2010.

Secondly, since not all relevant material is already digitised, extra financing has to come from projects funded by the EU and the national Scientific Research Organisations. An important EU-project is IMPACT, Improving Access to Texts (www.impact-project.eu), duration 2008 – 2011. Another project for collecting and digitising historical texts of the 15th and 16th Century has been submitted to the Dutch Organisation for Scientific Research in September 2009 (Notification of acceptance in March 2010).

Thirdly, the Taalbank Nederlands is exploring the possibilities of the web as corpus (WAC) to be seen as a dynamic corpus that could be very functional for the contemporary General Dictionary of Dutch, especially for the discovery of new words and new meanings or uses of existing words.

The third component of the Taalbank Nederlands, besides the dictionary component and the corpus component, is the lexicon component. This computational lexicon, named GiGaNT (Giant Integrated Lexicon of the Dutch Language), will contain all the entries and subentries, inclusive multiword expressions and named entities of all the digitised texts and INL-dictionaries, and all the existing word forms in all their possible variants throughout the different periods of Dutch; all linked to their modern Dutch equivalent. Every entry will be saved together with morphosyntactic information. The lexicon will be corpus-based: that means that for every word information will be stored about its attestation, frequency, location and date. On a longer term, also lexical-semantic information will be added to the computational lexicon. The main goal of this action is to start working on a so-called diachronic wordnet, which will be based on meaning equivalents rather than on form equivalents.

Why is there a need for this lexicon? First, when mapped onto a text or corpus of texts, all the known or recognised words and word forms will be tagged or labelled in the texts. In doing so, the user will be able to for example gain direct access to the lexicon and the connected dictionaries when clicking the words in the text. Secondly, simultaneously with the first action, a list of words, word forms and spellings will be generated which are not recognised when comparing the new material with the existing lexical material. These lists will be taken up by the lexicographers specialised in the different periods of Dutch; the missing words will be added to the lexical information of the correct period and will be handled as new entries.

In this way both the lexicon and the dictionaries will be extended, depending on the findings in the new material, and, at the same time, the digitised texts will be grammatically annotated, so that they can be used for further linguistic research. A very welcomed side-effect of this mechanism is that it will be possible on a short notice to click on a word in a text and look up its meaning or discover its meaning and changes thereof throughout the different periods of Dutch. Also the internet can in theory be used as such. When clicking on a word in the results googled, the users will get the corresponding linguistic information instantly.

4.1. A first U-turn: from manual labour and printed material to computational linguistics and the internet

Let me start this section by giving an overview of the personnel at the INL 15 years ago, say the mid 90s, and now. In the 90s the majority of the personnel were assistants, personnel that looked up the quotations manually, wrote everything down on paper filing systems which were delivered to the lexicographers to be treated accordingly. Each member of the personnel was attached to only one dictionary. The department of Automatisation consisted of 1 computational linguist, 1 system analyst and 2 programmers, on a total of about 50 personnel. Nowadays there still are about 50 personnel at the INL. The department of Automatisation consists now of 3 to 4 system analysts, 3 software engineers and 8 computational linguists. The amount of assistants has decreased enormously. Besides that, each member of the personnel functions within a pooling system, except for the lexicographers who are still connected to one or more projects concerning the highly specialised descriptive work which is a conditio sine qua non for the historical dictionaries.

The server park consisted in the 90s of a few minor systems together with some desktops. Nowadays an investment of about 800.000 euro has been made to increase the storing capacity above 9 terabyte amongst other improvements; every member of the personnel has at least one desktop and there are three technical environments, one for developing, another for testing and another for production. Looking at the software, nowadays the computational

linguists together with the software engineers have been developing dictionary writing systems (DWSs) containing data and content management systems and have been focusing on the functional and graphical designs or user interfaces for the web, to support the users when consulting the corpora and dictionaries. Furthermore, every dictionary has to be digitised corpus-based and communities are to be created in order to establish a daily interaction between the users of the dictionaries and the developers, being the lexicographers and computational linguists. Updates of the dictionaries are called releases; they have to be planned in at least once or twice a year. Recently a service desk has been developed using Quism enabling both users and developers to intensify their interaction.

4.2. A second U-turn: from single functionality to multiple functionality
In the 1990s, dictionaries were structured alphabetically, developed alphabetically and had one major function: to give the meaning of a word or a multi-word expression. Nowadays, dictionaries are developed as semantic modules. That means that the slots and fillers are dealt with in a more horizontal way, completing the pattern, instead of working vertically, i.e. providing all different types of information connected with one entry at once. More information is given about the morphological composition of the words and their syntactic behaviour. The users are no longer satisfied with the traditional semasiological function for single words, they also want to look up multi-word expressions, collocations and idioms, and, the most important change, they want to consult the dictionaries in an onomasiological way, i.e. giving a short description of the meaning of a word and receiving one or more forms as the output of the search procedure. Taking all of the aforementioned into account, one of the most important key words for the lexicographers has become consistency, especially with regard to the fillers that make concrete each slot or type of information.

The information should no longer be restricted to written language. Spoken language is required for the correct pronunciation of the words, sounds are required for illustrating e.g. the typical sounds of animals; films and 3d-material is required to illustrate e.g. action verbs or the parts of a vehicle. These multimedia demands are not only put upon contemporary dictionaries but also on historical dictionaries, e.g. to illustrate concepts or artefacts that are no longer known. All in all, the traditional explanatory dictionaries are bit by bit replaced by linguistic compendia or thesauri containing encyclopaedic multimedia information.

To put it short: the dictionaries and lexicons have become the central treasures in which all types of information can be found, be it pure linguistic information or encyclopaedic information, using different types of questions, be it semasiological, onamasiological or numerical (e.g. how many, how often etc.), and in which the information is presented by different means, be it written, oral (sounds) or visual (graphical).

4.3. A third U-turn: from stand-alone product to spin-offs, linking and integration
It is a fact that in the future it will become more and more difficult to find financing for changing the basic databases and their contents. Money however can be found for minor maintenance and for filling the time gaps. Minor investments can also be expected for correction and minor updates such as what can be called tick-box lexicography, the cutting and pasting of examples out of more recent texts replacing older ones and the detection of neologisms and new meanings of existing words.

The INL-dictionaries and lexicons can be seen as the core material from which other products can be derived by third parties. Since the Institute is financed by public funding, tax payer money, the free use of its products must be guaranteed for research, education and the

interested every day users. Third parties with an economic interest however have to pay an approved market price as they do for other products that are subsidised by tax payer money. Third parties are also interested in connecting or linking their data to the INL-data in order to make those data more accessible to the general public.

Finally, some other parties seek a full integration with the INL-data. The *Etymologiewoordeboek van Afrikaans* (EWA, Etymological Dictionary of Afrikaans) has been fully integrated in the online application of the *Woordenboek der Nederlandsche Taal*; the online version of the *Woordenboek der Friese Taal* (WFT, the Dictionary of the Frisian Language) has - as you probably all know - been presented recently.

A very important second type of linking and integration concerns the relation between the dictionaries and the text or corpus material. As said before, the IMPACT project tries to improve the access to historical documents by using language models and dictionaries, complementing the OCR. Important to notice is that within this project the linguistic processing in the OCR engine and the postprocessing will be adapted to historical language by using historical lexica and models of historical spelling.

As mentioned before, the connection between two types of data, i.e. lexical items and text data, will point at the gaps that still exist in the dictionary material. Not only the linking of historical texts to historical dictionaries seems to be important enough to be funded by the EU, also the linking of present day dictionaries to the massive volume of texts on the web is seen as a topic for which the necessary funding can be found. The Institute will work intensively on exploring this possibility from 2010 onwards.

5. Quo vadis lexicography?

From 2011 onwards - that is after ending the chapter of the *Woordenboek der Nederlandsche Taal* - the Institute will consist of three departments that will take care of the production and enclosure of corpora, dictionaries and lexicons. The three departments are, firstly, the Taalbank Nederlands, secondly, the Department of Description and Production, now filled by the *Algemeen Nederlands Woordenboek*, and, thirdly, the TST-Centrale (HLT Agency) which I left out of this presentation altogether.

There will still be of course also the supportive Department of Automatisation which I didn't elaborate on here either. The Taalbank Nederlands will focus on its four major tasks, i.e. acquiring and annotating corpora (historical texts and the web), dictionaries (maintaining, correcting, updating and linking the INL-material and the material provided by third parties), the linking between corpora and dictionaries, and the GiGaNT lexicon in which all of the linguistic descriptive information from 500 onwards will be stored, maintained, updated and integrated. An important issue will be the user friendliness on the web. Hereby is meant the different ways the linguistic, no longer just the lexical, information will be made available for all types of users. Special attention will be paid to elaborate the onomasiological search and the linking of all historical entries to their modern equivalents. All of the INL-material will furthermore be linked and integrated with historical text or corpus data from third parties. Also, the level of recognition of scanning which has been lifted from the character level to the word level, could be broadened towards multi-word expressions, collocations and idioms. Finally, the INL-data will be connected to data of other reference databases, for example the different dictionaries of the dialects and of related languages, e.g. Frisian or Afrikaans.

The other department is concerned with description and production. From 2011 onwards, all of the historical dictionaries of Dutch will be integrated and they will be available online. The remaining descriptive work by then is the development and expansion of the *Algemeen Nederlands Woordenboek* (ANW, Dictionary of Contemporary Dutch). While the crucial words for the historical dictionaries are linking and integration, the most important issue for the contemporary dictionary is integration and extension. Extension with different media, e.g. sounds and graphics, extension with different types of linguistic information, e.g. morphosyntactic information and the linking with reference databases of Dutch phonology, morphology and syntax. In fact a proposal was submitted to the Dutch Scientific Foundation to finance a project called The Language Portal of Dutch, asking for about 3.5 million euro for the development of a digitised grammar of Dutch which will be linked to the INL's *Algemeen Nederlands Woordenboek* (Notification of acceptance in March 2010). The last and perhaps most difficult, but also the most challenging extension, however, concerns the onomasiological search within all of the online dictionaries that are available at the INL.

In short: lexicography in its broadest sense has a great future ahead at the Instituut voor Nederlandse Lexicologie!

Jeannine Beeken

References

Algemeen Nederlands Woordenboek (ANW, Dictionary of Contemporary Dutch). http://anw.inl.nl
Etymologisch Woordenboek van het Nederlands (EWN, Etymological Dictionary of Dutch). http://www.etymologie.nl/
Middelnederlandsch Woordenboek (MNW, Dictionary of Middle Dutch). http://mnw.inl.nl
Oudnederlands Woordenboek (ONW, Dictionary of Old Dutch). http://onw.inl.nl
Vroegmiddelnederlands Woordenboek (VMNW, Dictionary of Early Middle Dutch). http://vmnw.inl.nl
Woordenboek der Nederlandsche Taal (WNT, Dictionary of the Dutch Language). http://wnt.inl.nl