Dictionary Use and Language Games:
Getting to Know the Dictionary as Part of the Game

Tanneke Schoonheim, Carole Tiberius, Jan Niestadt & Rob Tempelaars

Abstract

Most electronic dictionaries promise dynamic, proactive search via multiple criteria and via diverse access routes, but, often, they do not realise their full potential and their search options are still limited to the traditional search from word to meaning. The ANW (Algemeen Nederlands Woordenboek) - a free online scholarly dictionary of contemporary standard Dutch, which is currently being compiled at the Instituut voor Nederlandse Lexicologie (INL) - is different. It offers a range of search strategies, helping the user both with encoding and decoding tasks.

In December 2009, a demo version of the dictionary was launched. The dictionary is updated on a regular basis with an average of 500 to 750 new entries each time. An analysis of the log files shows that since its launch the average use of the dictionary is fairly stable, except for November 2010, when it almost tripled as a result of a language game, Het Verloren Woord ('The Lost Word') that INL launched. During a period of 6 weeks, participants received every week one or more cryptic descriptions or instructions in order to find the 'lost' word. Each description and/or instruction gave part of the word away and after solving all cryptic descriptions, the lost word, could be found in the ANW. The game attracted almost 2,000 players, who for several weeks explored the ANW thoroughly, using all the search facilities that are offered.

We will discuss the effect of this language game on the use of the ANW dictionary. In addition, we will show how a language game can play an educational role in familiarising users with the new possibilities that online dictionaries offer.

1. Introduction

In this paper, we will discuss the effect of a language game on the use of a dictionary, the ANW (*Algemeen Nederlands Woordenboek*). The ANW is a free online scholarly dictionary of contemporary standard Dutch of which a demo version was launched in December 2009. An analysis of the log files shows that since its launch the average use of the dictionary is fairly stable, except for November 2010, when it almost tripled. In November 2010, the Institute for Dutch Lexicology (INL) launched a language game, Het Verloren Woord ('The Lost Word'). This game attracted almost 2,000 players, who for several weeks, as our log files show, explored the ANW and the other dictionaries of INL thoroughly, using the different search facilities that are offered. We will conclude this paper by showing how a language game such as Het Verloren Woord can play an educational role in familiarising users with the new possibilities that online dictionaries offer.

2. The ANW dictionary

The ANW (http://anw.inl.nl) is a comprehensive online scholarly dictionary of contemporary standard Dutch in the Netherlands and in Flanders, the Dutch speaking part of Belgium (Moerdijk 2004, 2008; Schoonheim & Tempelaars 2010). The dictionary focuses on written Dutch and covers the period from 1970 onwards. It contains a detailed lexical-semantic description of the core vocabulary of contemporary Dutch, as well as a representative selection of neologisms from the year 2000 onwards. Editing and publication of the lemmas is

done modularly, mainly on the basis of their semantic class, which increases efficiency and leads to greater consistency throughout the dictionary.

The dictionary was conceived as an online dictionary right from the outset and aims to exploit the online medium to the full offering the user a range of search possibilities supporting both semasiological and onomasiological queries. Four search options are distinguished (Moerdijk, Tiberius & Niestadt 2008; Tiberius & Niestadt 2010):

a) **Word to Meaning**, i.e. search for information about a word or phrase;
b) **Meaning to Word,** i.e. search for a word starting from the meaning;
c) **Features to Words**, i.e. a detailed search for words sharing one or more common features;
d) **Examples**, i.e. search for example sentences.

The idea is that by presenting the search options in this way, users have a better overview of what they can actually search for and will be more enticed to explore the various search options that are new to them. To see whether our hypothesis was confirmed an analysis of the log files was carried out after the first 20 months of use.

3. The Language Game: Het Verloren Woord

Het Verloren Woord ('The Lost Word') was an online language game developed by INL and made available to users during a six-week period from 1 November till 15 December 2010 (see: http://www.hetverlorenwoord.eu). The purpose of the game was to familiarise a broader public with the dictionaries of the institute.

The game started with an introductory film, in which one Dutch and one Belgian TV personality explained that the Dutch language had lost a word and made an appeal to the public for help to find it. In total, 2355 people watched the introductory film on YouTube (www.youtube.com/watch?v=jHXblpPb5MA). Of those, 1883 registered for the game and after six weeks of puzzling, 70 participants sent in a solution, of which 66 were correct.

In order to find the lost word, participants had to solve a set of cryptic descriptions which they received over a period of six weeks. An example of such a cryptic description was given in the introductory film:

Round 1: *Verlang voor dit gevecht naar het omgekeerde van niet vroeg.*
'Long for this fight for the reverse of not early.'

Solution: *taalstrijd* ('linguistic conflict').
Explanation: *niet vroeg* ('not early') = *laat* ('late'); the reverse of *laat* = *taal* ('language');
Extra clue: a synonym of *verlangen* ('long') = *talen*, with the imperative form: *taal*. Put *taal* before a word for *gevecht* ('fight'); another word for *gevecht* = *strijd*. Thus: *taalstrijd* ('linguistic conflict').

The cryptic descriptions were released on Monday at a set time. On Thursday, also at a set time, a clue was given to help to solve the cryptic description. All information was given to

the players and followers of the game via email and Twitter. Twitter was used as a communication medium to ensure that players would not only exchange ideas with the organisers of the game but also amongst each other. For this a forum was created on the website were players could leave their comments and findings. In addition a second forum was created on a popular website for cryptograms in the Netherlands and Flanders (www.jasperscryptogrammensite.com) where players could comment on the game, discuss the cryptic descriptions and help each other with clues without revealing the solutions.

To solve these weekly puzzles, participants had to use the dictionaries of INL which are all available online, i.e. *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) and the *Algemeen Nederlands Woordenboek* (ANW, 1970 –). Each week, the participants could send their answers to the organisers of the game who would indicate whether the solution was correct or not. Correct answers would result in a number of letters which, after six weeks, would form two words leading to the 'lost' word.

So, after five weeks, successful participants would have found the word parts *ra*, *er*, *baar*, *cu*, *lijk*, *wond*, *mi* en *leus*, enabling them to make the synonyms *wonderbaarlijk* and *miraculeus* both meaning 'miraculous'. With this information they entered round 6, the round in which they were to find the 'lost' word at last: *taalobstakel* ('language obstacle').

4. Getting to know the dictionary as part of the Game

4.1. *Analysis of log files in general*

In December 2009, a demo version of the ANW dictionary was launched containing 914 lemmas. The dictionary is updated on a regular basis (every three months) and new entries are being added each time. In July 2011, an evaluation was carried out on the basis of a detailed quantitative analysis of the log files. Here we limit our discussion to those parts of the analysis which are relevant to our discussion of the effect of the language game on the log files.

During the first 20 months, the ANW server served 650,000 pages, performed 120,000 searches, from 180,000 different IP addresses, in 240,000 sessions. During the first few months after its launch the dictionary is not used that much, but an increase can be observed after six months with a clear peak in November 2010. Thereafter the average use seems to stabilise at around 40,000 requests per month.
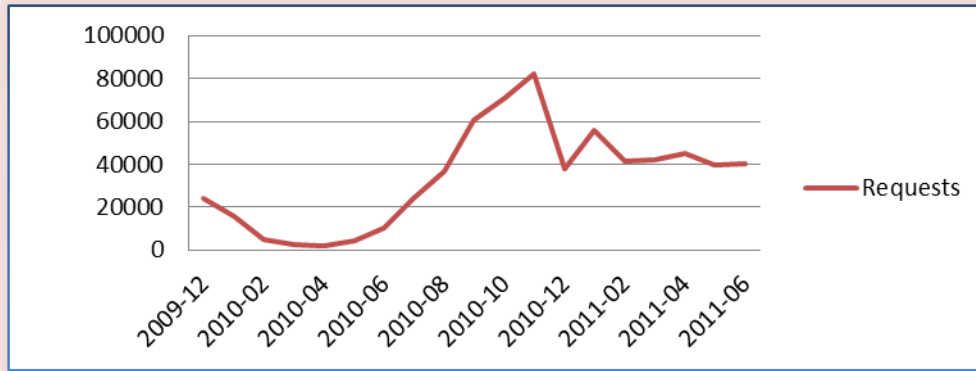
Figure 1. Number of requests to the ANW dictionary per month.

If we look at the kind of queries users ask, we see that the main use of the dictionary is to look up the meaning of a word. This is maybe not so surprising as this is the traditional use of a dictionary.
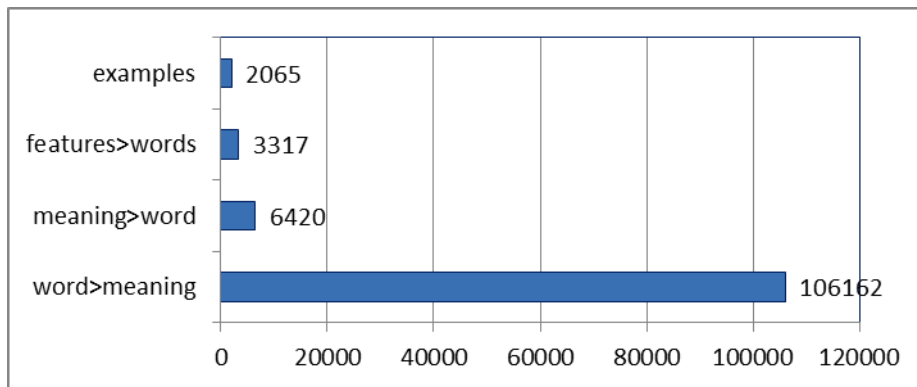

Figure 2. Number of queries per type.

Unfortunately, the other search options do not get used as much as we had hoped for, but, we do find it encouraging seeing that users are starting to explore these search options, which are new to them in a dictionary which is still under construction.

Furthermore, an analysis of the content of the queries found under the search option **Word to Meaning** shows that there are actually quite a few instances where, in theory, the query belongs to one of the other search types, e.g.

- **Features to Words**
  - Woorden met … *(words with ...)*
  - Soort … *(Kind of ...)*
- **Meaning to Word**
  - vorm van pasta *(kind of pasta)*
  - vrouw uit Frankrijk *(woman from France)*
  - tas waar je al je pennen in kan doen *(bag in which you can put your pens)*

This may be an indication that users would prefer to have one search box for all their queries. Further discussion of this goes, however, beyond the scope of this paper.

977

4.2. *Analysis of log files Het Verloren Woord*

Right from the start of the game Het Verloren Woord, an increase in the number of visitors of the INL website could be observed which is also reflected in the ANW log files. To establish whether there is a link between the language game and the use of the ANW, a content analysis was carried out of the most frequent queries for each of the search options offered by the ANW dictionary. Looking at the results, we observed that searches related to the 'Lost Word' are omnipresent in each of the search types offered by the ANW.

Let us start with the most popular search type, i.e. **Word to Meaning**. Looking at the ten most frequent queries, we see that there are three related to the language game, i.e. *palindroom* ('palindrome'*)* being the second most frequently looked up word, *wonderbaarlijk* ('miraculous') being the seventh most frequent query and *taalobstakel* ('language obstacle') being the ninth most frequent query. An analysis of the first 500 searches shows that a further six words related to Het Verloren Woord can be added to this list:

palindroom (palindrome)	93 (2)
wonderbaarlijk (miraculous)	66 (7)
taalobstakel (language obstacle)	64 (9)
giftig (poisonous)	52 (18)
miraculeus (miraculous)	52 (19)
obstakel (obstacle)	47 (23)
hindernis (obstacle)	37 (52)
sovjetgabber ('sovjet pal')	15 (432)
verstopt (hidden)	15 (438)

The numbers in brackets indicate their rank in the frequency list.

The reverse search (**Meaning to Word**) has also been used quite extensively in order to find the lost word. The six most frequent descriptions that are being looked up, are all related to the game: *hindernis, wonderbaarlijk, verstopt, miraculeus, taalobstakel, obstakel.*
It is interesting to note though that players generally did not attempt to run more complex queries in which they combined search terms.

Even when we look at the least used search option (**Features to Words**) we see familiar words of the game pop up:

Article contains the word:	giftig middeleeuws ('poisonous' 'medieval')
Definition contains the word:	giftig ('poisonous')
Singular is:	palindroom ('palindrome')

People also tried to find the answer to the cryptic descriptions using the example sentences that are cited in the ANW, searching frequently for example sentences containing words as *wonderbaarlijk*, *taalobstakel*, *miraculeus, hindernis, obstakel, verstopt, palindroom* and *giftig.*

5. Conclusion

The purpose of the game was to try to get the Institute for Dutch Lexicology (INL) as well as its dictionaries, both historical and contemporary, better known. By guiding the participants

by means of cryptic descriptions through the dictionaries, they would become more familiar with their structure and content, as well as with the additional search options that they offer. This latter goal was more or less achieved, but the degree of difficulty of the game turned out to be too high to achieve the first goal. Only 70 participants sent in an answer after six weeks and remained involved in the game till the end.

Feedback of these participants reveals that overall they were very pleased with the game apart from some initial difficulties. The challenge of solving the cryptic descriptions, the possibilities offered by the INL dictionaries and the service of the organisers were highly appreciated and they can hardly wait for a follow-up.

The log files show that the participants have indeed been introduced to aspects of dictionaries that were completely new to them, such as the possibility to search for a word on the basis of its meaning or definition or to search for words combining one or more features which are both available in the ANW. At least they know now about the INL dictionaries and what information can be found in them and which search strategies can be used.

However, the game did not fully achieve its goal of getting the institute and its dictionaries better known by the general public as at the end of the game usage of the dictionaries dropped again. Also, the game was too difficult and not enough people participated. If INL is to develop another language game, this will be taken into account, and a larger audience will be targeted.

References

Moerdijk, F. 2004. 'Het Algemeen Nederlands Woordenboek (ANW).' *Nederlandse Taalkunde* 9: 175–182. http://taalkunde.letterentijdschriften.nl/document_articles/208.pdf.

Moerdijk, F. 2008. 'Frames and Semagrams. Meaning Description in the General Dutch Dictionary.' In E. Bernal and J. DeCesaris (eds.), *Proceedings of the XIII Euralex International Congress: Barcelona, 15-19 July 2008*. Barcelona: L'Institut Universitari de Lingüistica Aplicada, Universitat Pompeu Fabra, 561–571.

Moerdijk, F., C. Tiberius and J. Niestadt 2008. 'Accessing the ANW Dictionary.' In *Proceedings of the Workshop on Cognitive Aspects of the Lexicon (COGALEX 2008)*. Coling 2008 Organizing Committee, 18-24. www.aclweb.org/anthology/W08-1903.

Schoonheim, T. and R. Tempelaars 2010. 'Dutch Lexicography in Progress, The Algemeen Nederlands Woordenboek (ANW).' In A. Dykstra and T. Schoonheim (eds.), *Proceedings of the XIV Euralex International Congress*, *Leeuwarden, 6-10 July 2010*. Ljouwert: Fryske Akademy / Afuk.

Tiberius, C. and J. Niestadt 2010. 'The ANW: an online Dutch Dictionary.' In A. Dykstra and T. Schoonheim (eds.), *Proceedings of the XIV Euralex International Congress*, *Leeuwarden, 6-10 July 2010*. Ljouwert: Fryske Akademy / Afuk.