



INSTITUT FÜR
DEUTSCHE SPRACHE

OPAL

Online publizierte Arbeiten zur Linguistik

ISSN 1860-9422

Sonderheft

1/2008

Fons Moerdijk

Das „Algemeen Nederlandse Woordenboek“ (ANW) und *elexiko* – ein Vergleich

aus: Klosa (Hg.): Lexikografische Portale im Internet.

(= OPAL-Sonderheft 1/2008), S. 143-151.

OPAL – Online publizierte Arbeiten zur Linguistik
Herausgegeben vom Institut für Deutsche Sprache



Institut für Deutsche Sprache
Postfach 10 16 21
68016 Mannheim
opal@ids-mannheim.de

Technische Redaktion: Norbert Volz

© 2008 IDS Mannheim – Alle Rechte vorbehalten

Das Werk einschließlich seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechts ist ohne Zustimmung der Copyright-Inhaber unzulässig und strafbar. Das zulässige Zitieren kleinerer Teile in einem eigenen selbstständigen Werk (§ 51 UrhG) erfordert stets die Angabe der Quelle (§ 63 UrhG) in einer geeigneten Form (§ 13 UrhG). Eine Verletzung des Urheberrechts kann Rechtsfolgen nach sich ziehen (§ 97 UrhG). Dies gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen. Die zugänglichen Daten dürfen von den Nutzern also nur zu rein wissenschaftlichen Zwecken genutzt werden. Eine darüber hinausgehende Nutzung, gleich welcher Art, oder die Verarbeitung und Bearbeitung dieser Daten mit dem Zweck, sie anschließend selbst oder durch Dritte kommerziell zu nutzen, bedarf einer besonderen Genehmigung des IDS (Lizenz). Es ist nicht gestattet, Kopien der Textdateien auf externen Webservern zur Verfügung zu stellen oder Dritten auf sonstigem Wege zugänglich zu machen. Bei der Veröffentlichung von Forschungsergebnissen, in denen OPAL-Publikationen zitiert werden, bitten die Autoren und Herausgeber um eine entsprechende kollegiale Information an opal@ids-mannheim.de.

Fons Moerdijk

Das „Algemeen Nederlandse Woordenboek“ (ANW) und *ellexiko* – ein Vergleich

Abstract

ANW und *ellexiko* repräsentieren eine neue Generation von wissenschaftlichen elektronischen (Online-) Wörterbüchern: sie sind keine digitalisierten Klone von schon existierenden Printwörterbüchern, sondern werden inhaltlich neu und mit voller Berücksichtigung der Möglichkeiten des neuen Mediums realisiert. In diesem Beitrag werden zuerst pauschal einige wichtige Parallelen und Unterschiede zwischen dem ANW und *ellexiko* beleuchtet. Anschließend wird der substanzielle Unterschied in den Suchoptionen eingehend behandelt. *ellexiko* hantiert mit dem Unterschied „einfache Suche“ neben „Expertensuche“ – ein bekanntes System. Das ANW hat ein eigenes, neues System mit den folgenden Suchmöglichkeiten entwickelt: Suche nach Information zu einem Wort, Suche nach einem Wort (von der Bedeutung aus), Suche nach Wörtern (auf Grund eines oder mehrerer gemeinschaftlicher Merkmale), Suche nach Beispielen mit gemeinsamen Merkmalen und Suche nach Information über das Wörterbuch selbst. In den onomasiologischen Suchformen, die vom Inhalt zum Wort führen, spielt das „Semagramm“, die Darstellung von Kenntnis, die mit einem Wort zu verbinden ist, in einem Rahmen mit „Slots“ und „Fillern“ eine substanzielle Rolle. Das Semagramm ist eine weitere Erneuerung des ANW.

The ANW and *ellexiko* represent a new generation of scholarly online dictionaries: they are no digitalised copies of existing printed dictionaries, but they are, as far as content and technology is concerned, developed from scratch as brand-new products. First a short overview is given of the main similarities and differences between the two dictionaries. Then the differences in search options offered by *ellexiko* and the ANW are discussed in more depth. *ellexiko* follows the well-known system with the distinction ‘simple search’ and ‘advanced search’. The ANW has created a new system with the following search options: search for information about a word, search for a word (with the meaning as starting-point), search for words (on the basis of one or more common features), search for examples (with common features) and search for meta-information about the dictionary itself. In the onomasiological searches from content to word the ‘semagram’, the representation of conceptual knowledge connected with a word in a frame with ‘slots’ and ‘fillers’, plays a prominent role. The semagram is also an innovation of the ANW.

Inhalt:

1. Parallelen und Unterschiede zwischen ANW und *ellexiko*
 - 1.1 Parallelen
 - 1.2 Unterschiede
2. Suchformen im Allgemeinen
 - 2.1 Das Dichotomiesystem
 - 2.2 Einwände gegen das Dichotomiesystem
3. Spezielle Suchformen bei *ellexiko* und im ANW
 - 3.1 Die *ellexiko*-Suche
 - 3.2 Die ANW-Suche
4. Abschluss
5. Literatur
6. Nachschlagewerke

Das *elexiko*-Wörterbuch,¹ ein Produkt des Instituts für Deutsche Sprache, ist schon einige Jahre im Internet zu bewundern. Etwa 800 Wortartikel zu hochfrequenten Stichwörtern sind bereits redaktionell bearbeitet und publiziert. Es bedarf also keiner näheren Einführung. Das gilt nicht für das „Algemeen Nederlandse Woordenboek“ (ANW). Das ANW wird ein Online-Wörterbuch der niederländischen Gegenwartssprache und ist ein Projekt des Instituts für Niederländische Lexikologie (INL) in Leiden. Im Gegensatz zu *elexiko* ist es noch nicht an die Öffentlichkeit gebracht. Erst mit Anfang des Jahres 2009 sind die ersten Resultate im Internet zu erwarten.

Mit dem ANW wurde im Jahre 2001 begonnen. Der Zeitraum von 2001 bis 2006 stand im Zeichen zweier Hauptaufgaben: Korpusaufbau und Konzeptentwicklung. 2006 wurden Probeartikel verfasst. Sehr viel Zeit hat man der inhaltlichen Analyse von Definitionen aus anderen Wörterbüchern gewidmet, auf deren Grundlage der Wortschatz in Bedeutungsklassen unterteilt wurde. Aus diesen jeweiligen Klassen wurden Schablonen für die semantische Beschreibung von Wörtern aufgestellt.

In diesem Beitrag behandle ich zuerst, in sehr groben Zügen, einige wichtige Parallelen und Unterschiede zwischen dem ANW und *elexiko*. Anschließend werde ich näher auf einen sehr substanziellen Unterschied, der mit der Präsentation im Internet zusammenhängt, und zwar den Unterschied bei den Suchoptionen, eingehen. Ich habe mich dafür entschieden, weil das ANW sich in dieser Hinsicht nicht nur von *elexiko*, sondern meines Wissens auch von anderen wissenschaftlichen Online-Wörterbüchern unterscheiden wird.

1. Parallelen und Unterschiede zwischen ANW und *elexiko*

1.1 Parallelen

ANW und *elexiko* repräsentieren eine neue Generation von wissenschaftlichen elektronischen (Online-)Wörterbüchern. Ich spreche von einer neuen Generation; erstens, weil sie keine digitalisierten Klone von schon existierenden Printwörterbüchern sind, sondern inhaltlich einen ganz neuen Anfang machen; und zweitens, weil sie von Anfang an in ihrem Entwurf die speziellen neuen Möglichkeiten und Perspektiven des Wörterbuchs als Hypertext berücksichtigt haben (Moerdijk 2007a).

Beide Wörterbücher beschreiben die allgemeine (vorwiegend geschriebene) Gegenwartssprache (unter Berücksichtigung der nationalen Sprachvarietäten, der „Natiolekte“). Sie wollen dabei einem breiten Spektrum von Benutzern (von Laien bis zu Sprachwissenschaftlern) dienen. Wie *elexiko* widmet das ANW Neologismen spezielle Aufmerksamkeit.

Eine dritte Übereinstimmung ist, dass sie „korpusbasiert“ sind. Sie werden zusammengestellt mithilfe eines Korpus im modernen Sinne einer Sammlung elektronischer Texte, die auf repräsentative Weise die Sprachwirklichkeit, die man beschreiben will, widerspiegeln.

Weiterhin haben beide eine sehr umfassende, reichlich verzweigte Artikelstruktur. Das ist teils eine Konsequenz der vielgestaltigen Zielgruppen, teils bedingt durch Bedürfnis-

¹ *elexiko* wird im Folgenden im Stand von Mai 2007 präsentiert, da sich am hier Beschriebenen inhaltlich nichts Grundsätzliches geändert hat.

se der Wissenschaftler, die Material für ihre Untersuchungen brauchen. Sie stellen vor allem Fragen, die von bestimmten Merkmalen aus zu den Wörtern führen, die diese Merkmale besitzen. Es sind gerade solche Fragen, die eine Ausbreitung der Suchmöglichkeiten und deshalb der Mikrostruktur erfordern. Für die Festlegung dieser Struktur benutzen sie die Markierungssprache XML. Jeder Wörterbuchartikel ist ein XML-Dokument und wird als solches in einer Datenbank gespeichert.

1.2 Unterschiede

Obwohl beide Wörterbücher ein Korpus als empirische Basis haben, gibt es doch auch einen wichtigen Unterschied. *ellexiko* konnte sich eines bestehenden Korpus bedienen: des Deutschen Referenzkorpus des IDS Mannheim. Das *ellexiko*-Korpus umfasst derzeit ca. 1,3 Milliarden Textwörter und besteht aus Zeitungs- und Zeitschriftentexten. Das ANW hat sich sein Korpus selbst bauen müssen. Das hat einen großen Teil der Vorbereitungsperiode beansprucht. Das ANW-Korpus hat einen Umfang von 104 Millionen Textwörtern. Wir haben gezielt eine Diversität von Themen und Genres angestrebt. Für etwa 1500 verschiedene Gesellschaftsgebiete wurden repräsentative Texte aus dem Internet heruntergeladen. Auch ein Teilkorpus aus literarischen Texten wurde zusammengestellt. Nur ungefähr 40% der Korpus-Texte stammten aus Zeitungen und Zeitschriften.

Beide Wörterbücher benutzen, wie gesagt, die Markierungssprache XML, aber auch hier gibt es einen Unterschied. *ellexiko* gebraucht als XML-Editor *XMetal*, das ANW verwendet *Altova XMLSpy* und verfügt dabei über ein selbst entwickeltes, „lexikografenfreundliches“ Bearbeitungsformular.

Die Bedeutungsbeschreibung besteht bei *ellexiko* im Anschluss an die lexikografische Tradition aus einer Definition, aus einer erklärenden Paraphrase. Beim ANW ist sie zweiteilig und besteht aus Definition und so genanntem „Semagramm“. Das Semagramm ist eine Innovation des ANW. Ein Semagramm ist die Darstellung der Kenntnis, die mit einem Wort verbunden ist, in einem Rahmen mit „Slots“ und „Fillern“. „Slots“ sind abstrakte konzeptuelle Strukturelemente, die sich auf Merkmale und Relationen beziehen, die für die Bedeutungsklasse, wozu ein Wort gehört, gelten. Sie werden mit konkreten Daten, die für das betreffende Wort gelten (die „Filler“), ausgefüllt. Ich komme später noch darauf zu sprechen (siehe auch Moerdijk 2007b).

Beim ANW findet die Bearbeitung nicht alphabetisch und nach Frequenzgrad statt, sondern nach so genannten „lexikografischen Typen“, Wörtern, die wegen gemeinsamer syntaktischer, morphologischer und semantischer Eigenschaften eine eigenständige Gruppe ausmachen. Ein solches Verfahren sollte zu besseren, einheitlicheren und systematischeren lexikografischen Beschreibungen und Erklärungen führen.

2. Suchformen im Allgemeinen

2.1 Das Dichotomiesystem

Ein wesentlicher Vorteil des elektronischen Wörterbuchs ist, dass es verschiedene Zugriffswege zu den Wörterbuchdaten bietet, während das prototypische Wörterbuch nur den alphabetischen Weg über die Stichwörter zu bieten hat. Im alphabetischen Druckwörterbuch geht man von der Form zum Inhalt. Das elektronische Wörterbuch besitzt diese Möglichkeit auch, darüber hinaus bietet es mehrere Suchmöglichkeiten, die

von inhaltlichen Merkmalen zu den Wortformen führen. Mit einem bekannten terminologischen Unterschied aus der traditionellen Lexikologie kann man hier von semasiologischen (von Form zu Inhalt) und onomasiologischen Suchoptionen (von Inhalt zu Form) sprechen.

Im Allgemeinen werden die verschiedenen Suchmöglichkeiten von den Wörterbüchern selbst in Dichotomien wie „basiszoeken“ gegenüber „uitgebreid zoeken“ (i-WNT), „simple search“ gegenüber „advanced search“ (i-OED), „recherche d'un mot“ versus „recherche assistée“ und „recherche complexe“ (i-TLF), „Standardmodus“ gegenüber „Expertenmodus“ oder „erweiterte Suche“ (GWDS, auf CD-ROM) dargeboten. Wir benennen dieses System in der Folge der Einfachheit halber als „das Dichotomiesystem“.

Die einfache Suche ist die Stichwortsuche, die Suche, die zum Artikel führt. Wenn wir von den Unterschieden in der Zahl der gebotenen Möglichkeiten und in der Präsentationsweise absehen, dann können wir sagen, dass es sich im Allgemeinen bei der ausführlichen, fortgeschrittenen Suche um eine Ausweitung der Suchoptionen und um die Möglichkeit von komplexeren Suchanfragen handelt, wie die Suche mit mehreren Suchtermini, die Suche in bestimmten Kategorien oder Kombinationen von Kategorien, die Suche mit Operatoren, die Auswahl des Kontextbereiches und Ähnliches.

2.2 Einwände gegen das Dichotomiesystem

Ein großer Nachteil des Dichotomiesystems ist, dass es alles andere als benutzerfreundlich ist. Die Benutzer wissen gar nicht, welche Optionen sich hinter Begriffen wie „einfache Suche“, „basiszoeken“, „simple search“, „komplexe Suche“, „erweiterte Suche“ oder „Expertensuche“ verstecken, sondern finden dies erst mit zunehmender Erfahrung heraus.

Und selbst wenn man sich einmal auf den betreffenden Webseiten befindet, ist häufig überhaupt nicht klar, was man nun mit all den angebotenen Auswahl Fenstern, Menüleisten und Buttons anfangen soll, um die erwünschten Suchergebnisse zu bekommen. Das ist insbesondere der Fall bei komplexen Suchanfragen, z.B. der Expertensuche (advanced search), wo die Benutzer ihre Suchanfrage durch das Ausfüllen von Eingabefeldern und das Anklicken von Menüpunkten in mehreren Auswahl Fenstern selbst zusammensetzen müssen. In vielen Fällen ist der sprachliche und computertechnische Jargon dabei ein zusätzliches Problem. Selbstverständlich gibt es Hilfetexte. Wer das Internet konsultiert, will aber schnell, direkt und deutlich informiert werden und vermeidet Aktionen, die dem im Wege stehen.

Ein weiteres Problem ist, dass (auch in den Hilfetexten) die Formulierungen der Suchoptionen häufig nicht mit den Suchanfragen, wie die Benutzer sie tatsächlich im Kopf haben, übereinstimmen. Diese Unzulänglichkeit wird verursacht durch die Vermischung der Suche des Computers mit der Suche der Benutzer. Ich gebe ein Beispiel: Bei dem iWNT hat man eine Suchoption „zoek woord in betekenis“. Sie ist vergleichbar mit der Suchfunktion „simple search on definitions“ im OED. Im Hilfetext wird das Ziel dieser Suchfunktion mit dem Beispiel „to search for the word *generous* in definition text“ näher erklärt. Die Benutzer suchen aber gar nicht nach dem Wort *generous* in allen Definitionen oder nach allen Definitionen, die das Wort *generous* enthalten. Sie suchen jedoch nach allen Wörtern, die Synonyme von *generous* sind (oder die in einer anderen bestimmten semantischen Relation mit *generous* stehen). Es ist der Computer, der in allen

„Definitionen“ des vollständigen Datenbestandes suchen soll, um diesen Suchwunsch der Benutzer zu erfüllen. Auch Suchfunktionalitäten wie „Suche im Volltext“ oder „Suche in Teilen des Artikels“ sind Formulierungen, die auf das Suchgebiet des Computers gerichtet sind und nicht auf dasjenige, was die Benutzer tatsächlich finden wollen.

Schließlich möchte ich bemerken, dass im Dichotomiesystem bei Weitem nicht alle Möglichkeiten des „umgekehrten“ Suchens, des Suchens von Merkmalen zu Wortformen, ausgeschöpft worden sind. Man hat sich mehr oder weniger beschränkt (besser: beschränken müssen) auf Suchoptionen, die auf einer Auswahl von Angabeklassen basieren, die mit einer begrenzten Anzahl von festen Angabewerten zu der Mikrostruktur gehören (zum Beispiel: Wortart, mit den Werten „Adjektiv“, „Adverb“, „Nomen“, „Verb“ usw.). Das Angebot wird eher durch praktische Verfügbarkeit als durch theoretische Überlegungen bestimmt. Im Prinzip können jedoch alle Kategorien der Mikrostruktur Grundlage für solche umgekehrten Suchanfragen sein.

3. Spezielle Suchformen bei *ellexiko* und im ANW

3.1 Die *ellexiko*-Suche

ellexiko unterscheidet zwischen der einfachen Suche nach einem Stichwort, mit der man zum Artikel gelangt, und der „Expertensuche“ und gehört also zu den Wörterbüchern mit einem „Dichotomiesystem“. Bei dieser Expertensuche kann man nach allen Wörtern mit bestimmten Merkmalen in einem der folgenden Bereiche suchen: Orthografie, Wortart, Grammatik, Wortbildung, sinnverwandte Wörter und semantische Klasse. Bei diesen Angabeklassen kann man aus einer begrenzten, in einer Menüleiste präsentierten Reihe von möglichen Werten den zutreffenden Wert durch Anklicken auswählen. Bei Wortart werden „beliebig“, „Adjektiv“, „Adverb“, „Nomen“, „Verb“ als Werte angeboten, bei Wortbildung „beliebig“, „Komposita“, „explizite Derivate“, „Konversionen“, „Kurzwörter“ usw. „Beliebig“ ist immer der Defaultwert. Das ist der Wert, den die Benutzer anfangs im Auswahlfenster zu sehen bekommen. (Man würde den Benutzern meiner Meinung nach übrigens einen besseren Anhaltspunkt bieten, wenn man nicht „beliebig“, sondern einen der anderen, inhaltlich signifikanteren Wert als Anfangswert zeigen würde.)

Die Einwände, die wir im Allgemeinen gegen das Dichotomiesystem vorbrachten, treffen in gewissem Maße auch auf *ellexiko* zu. Benutzer werden sich auch hier fragen, was sie unter „Expertensuche“ zu „suchen haben“. Was den Umfang der dargestellten Suchoptionen betrifft, gilt, wie für die anderen Wörterbücher, dass nur ein sehr geringer Teil der potenziellen Funktionalitäten, wenigstens bis jetzt, realisiert ist. Die reiche Mikrostruktur des *ellexiko*-Wörterbuchs bietet sicherlich viel mehr Möglichkeiten für Suchanfragen, die auf Angabeklassen oder festen Angabewerten basieren.

Es gibt aber auch Verbesserungen. Die Präsentation ist unkompliziert und übersichtlich und die Kategorien und ihre Eingabefelder werden präsentiert mit der Überschrift: „Suche nach Stichwörtern mit bestimmten Merkmalen“. Diese Formulierung schließt ohne Zweifel schon besser an das an, was die Benutzer haben wollen (obwohl „Suche nach allen Wörtern mit bestimmten Merkmalen“ vielleicht noch besser gewesen wäre).

3.2 Die ANW-Suche

Das ANW wird sich nicht an diese offensichtlich schon etablierte Ordnung des Dichotomiesystems anschließen. Es will eine positive Lehre aus dessen Schwächen ziehen und strebt deshalb nach einem benutzerfreundlichen System, das folgende Bedingungen erfüllt:

- Die Benutzer sollten möglichst direkt und klar die Informationen bekommen, die sie sich wünschen.
- Die Suchformulare sollten in solcher Weise formuliert und präsentiert werden, dass die Benutzer gleich sehen, wo und wie sie ihre Suchanfragen stellen müssen.
- Die Formulierungen der Suchoptionen sollten nicht auf das abgestimmt sein, was der Computer suchen soll, sondern auf das, was die Benutzer suchen.
- Die Benutzer sollten möglichst wenig Fragen selbst konstruieren müssen.
- Die Suchmöglichkeiten sollten erheblich ausgeweitet werden, namentlich jene, die von Merkmalen zu Wörtern führen.

Dies hoffen wir zu erreichen, indem wir folgende Suchformen unterscheiden (die Begriffe „Wort“ und „Wörter“ sind hier großzügig zu betrachten; andere lexikalische Einheiten wie Phrasen, Sprichwörter und Namen sind mit eingeschlossen):

- 1) Suche nach Informationen über ein Wort (Wort → Artikel)
- 2) Suche nach einem Wort (von der Bedeutung aus) (Wort ← Bedeutung)
- 3) Suche nach Wörtern (aufgrund eines oder mehrerer gemeinschaftlicher Merkmale) (Wörter ← Merkmale)
- 4) Suche nach Beispielen/Beispielsätzen mit gemeinsamen Merkmalen (Beispielsätze ← Merkmale)
- 5) Suche nach Information über das Wörterbuch selbst.

Bei der **Suche nach Informationen über ein Wort** bekommt man durch Eingabe in ein Eingabefeld oder durch Anklicken eines Wortes in einer Stichwortliste Zugang zu dem Artikel, der dem entsprechenden Wort gewidmet ist. Es ist die Suchform, die mit der traditionellen Suche in dem alphabetischen, semasiologischen Printwörterbuch korrespondiert. Man bekommt hier Antwort auf Fragen vom Typ „Was bedeutet Wort x?“, „Wie schreibt man Wort x?“, „Wie spreche ich Wort x aus?“, „Kommt Wort x auch im Singular vor?“, „Was ist die Herkunft von Wort x?“ usw. Es sind *par excellence* Fragen, die das Publikum sich stellt. Von einer allgemeinen Anfangsseite aus kann man mit Hyperlinks die Webseiten der einzelnen Artikelsegmente erreichen. Daneben hat man die Möglichkeit, den Artikel als Ganzes anzufordern.

Die **Suche nach einem Wort** soll Antwort geben auf Fragen vom Typ „Wie heißt dies?“, „Wie nennt man dies?“, „Was ist der Name dafür?“. Man geht von der Bedeutung, von inhaltlichen Elementen, hin zum Wort. Es ist die onomasiologische Fragestellung im reinen, strengsten Sinne. Solche Fragen stellen sich im täglichen Leben, wenn Wörter uns nicht einfallen, und sind außerdem beliebte Quizfragen. Weil der Gang von

der Bedeutung zum Wort hin nur gut gelingt, wenn die Wortwahl in der Definition mit der Wortwahl der Benutzer zusammenfällt – d.h. in den bestehenden elektronischen Wörterbüchern bis jetzt fast nie –, haben wir in unserer Struktur auch einer neuen Art und Weise der Bedeutungsbeschreibung, dem „Semagramm“, einen Platz eingeräumt. Wie wir schon sagten, ist ein Semagramm die Darstellung von Kenntnissen, die mit einem Wort zu verbinden sind, in einem Rahmen mit „Slots“ und „Fillern“. „Slots“ sind abstrakte konzeptuelle Strukturelemente, die sich auf Merkmale und Relationen beziehen, die für die Bedeutungsklasse, zu der ein Wort gehört, gelten. Z.B. für die Klasse der Tiernamen, mit „Säugetier“, „Vogel“, „Fisch“, „Insekt“, „Reptil“, „Amphibie“ und „Weichtier“ als Subklassen, enthält die Klassenschablone folgende Slots:

SINNLICHE WAHRNEHMUNG

- ·Farbe
- ·Aussehen
- ·Umfang / Größe
- ·Form
- ·Geräusch, Schall, Laut

ZUSAMMENSETZUNG

- Teil
- Bau, Konstruktion

FUNKTION ODER ZIEL

- Funktion, Ziel, Bestimmung

ENTSTEHUNG

- Herkunft, Ursprung
- Fortpflanzung

RAUM, PLATZ

- Behausung
- Herkunft
- Platz, Ort (geografisch)

ZEIT

- Alter

ZUSTAND

- Zustand allgemein
- Stadium

EIGENSCHAFT / QUALITÄT

- Geschlecht
- unstoffliche Eigenschaft

HANDLUNG ODER WIRKUNG

- Bewegung
- Wirkung, Einwirkung
- Verwendung, Nutzung
- Behandlung oder Bearbeitung
- Verhalten

BETEILIGTE PERSONEN ODER SACHEN

- Produkt, Frucht

WERTSCHÄTZUNG UND BELANG

- Anerkennung

VERGLEICH

- Ähnlichkeit

BEZIEHUNG UND ZUSAMMENHANG

- Beziehung oder Relation

Nehmen wir an, dass jemand das Wort *koe* (‘Kuh’) nicht finden kann. Er weiß aber, dass es sich um ein Tier handelt, das auf dem Bauernhof lebt, öfters schwarz-weiß gescheckt ist (jedenfalls in Holland), auf der Wiese oder im Stall steht und Milch produziert. Der Benutzer kann dann solch eine Merkmalschablone durchgehen und bei jedem Merkmal sozusagen „gelenkt“ ausfüllen, was er weiß. Was er ausgefüllt hat, wird mit dem Semagramm gematcht, das der Lexikograph zusammengestellt hat und das wie folgt aussieht (hinter dem Zeichen # werden Schlüsselwörter aus den kurzen Sätzen wiederholt und mit möglichen Alternativen ergänzt, um das Retrieval zu verbessern; bei dem Artikel *koe*, hier im Beispiel als *Kuh* wiedergegeben, sehen die Benutzer nur die Sätzchen):

EINE KUH

OBERKATEGORIE:	ist ein Tier # Tier; Säugetier; Wiederkäuer
KATEGORIE:	ist ein Rind # Rind; Wiederkäuer
FARBE:	ist öfters schwarz-weiß gescheckt, aber auch braun-weiß, ganz schwarz, braun oder weiß # schwarz-weiß; schwarz-bunt; braun-weiß; rotbunt; bunt; schwarz; schwarz gescheckt; weiß; braun; gescheckt; rostbraun
UMFANG:	ist groß # groß

GERÄUSCH:	muht, brüllt, macht ein Geräusch, das wir mit einem tiefen, langgezogenen „Muh“ nachahmen # muht; brüllt; muhen; brüllen; muh
TEIL:	hat ein Euter, Hörner und vier Mägen: Pansen, Netzmagen, Blättermagen, Labmagen # Euter; Hörner; Mägen; Pansen; Netzmagen; Blättermagen; Labmagen; Käse-magen
BAU:	ist grobknochig, grobgliedrig, von derbem Gliederbau # grobknochig; grobgliedrig
FUNKTION:	gibt Milch und Fleisch # Milch; Fleisch; Rindfleisch; Milchproduktion; Fleischproduktion
PLATZ:	lebt auf einem Bauernhof; steht auf der Wiese und im Winter im Stall # Bauernhof; Wiese; Weide; Weideland; Stall
ALTER:	ist erwachsen, hat gewöhnlich schon einmal gekalbt # erwachsen; gekalbt
GESCHLECHT:	ist weiblich # weiblich; Weibchen; weibliches Rind
EIGENSCHAFT:	ist nützlich und zahm; wird betrachtet als ein freundliches, faules, träges, dummes, neugieriges, soziales Tier # nützlich; zahm; freundlich; träge; dumm; neugierig; sozial
VERHALTEN:	grast und kät wieder # grast; grasen; kät wieder; wiederkauen
BEHANDLUNG:	wird täglich gemolken; wird geschlachtet # melken; schlachten
PRODUKT:	gibt Milch und Fleisch # Milch; Fleisch
WERTSCHÄTZUNG:	ist nützlich # nützlich

Damit ist die Chance, dass das Wort *koe* auch tatsächlich vom Computer gefunden wird, beträchtlich größer geworden. In Wörterbüchern, die *Kuh* definieren als „erwachsenes weibliches Rind“, wäre dies nicht gelungen.

Die Suche nach Wörtern findet statt auf der Grundlage eines Merkmals oder einer Kombination mehrerer Merkmale, die die Wörter gemeinsam haben. Im Prinzip können alle Kategorien, Subkategorien und festen Angabewerte der Mikrostruktur als Ausgangspunkte für solche Fragen benutzt werden. Es handelt sich hier um Fragen oder Suchaufträge wie: Zeige mir alle Wörter mit *-iser-* oder *-izer-* in der Mitte (Orthografie/Wortform), zeige mir alle Wörter mit englischer Aussprache oder alle Wörter mit Endbetonung (Aussprache), zeige mir alle Substantive (Wortart), alle Komposita mit einem Imperativsatz als erstem Glied (Wortbildung), zeige mir alle Wörter für Fahrzeuge (Bedeutung/Begriffsklasse), alle Wörter aus dem Golfspiel (gesellschaftlicher Bereich), alle französischen Lehnwörter (Etymologie), alle Neologismen aus den Jahren 2000-2003 (Datierung/Etymologie) usw. Auch kombinierte Fragen sind möglich, zum Beispiel: Zeige mir alle Substantive, die in den Niederlanden einen Plural auf *-en*, aber in Belgien

einen Plural auf -s haben (Kategorien: Wortart, Substantiv Pluralform, Sprachvarietät). Diese Art von Fragen wird vor allem von Fachleuten, die Material für ihre Untersuchungen brauchen, gestellt. Onomasiologische Fragen sind hier die Fragen nach Benennungen für Begriffe oder nach Wörtern innerhalb bestimmter gesellschaftlicher Domänen. Insofern andere Fragen von Merkmalen zu Wortformen führen, könnte man auch von onomasiologischen Fragen sprechen, wobei dann onomasiologisch in weiterem (und nicht ganz exaktem) Sinn aufzufassen ist.

Die Suche nach Beispielsätzen bietet u.a. die Möglichkeit, Beispielsätze aus einer bestimmten Quelle zu sammeln, mit einem bestimmten Wort oder mit mehreren bestimmten Wörtern, von einem bestimmten Autor oder einer bestimmten Autorin, aus einem bestimmten Jahr oder einem bestimmten Zeitabschnitt. Auch hier sind Kombinationen möglich. Diese Fragen beziehen sich nicht auf Wörter und fallen deshalb aus der semasiologischen wie der onomasiologischen Ordnung. Es sind übrigens meiner Meinung nach Fragen, die man eher einem Korpus als einem Wörterbuch stellen sollte. Man kann aber das Wörterbuch als ein Korpus benutzen.

Bei der Suche nach Informationen über das Wörterbuch selbst geht es um Fragen wie: Was für ein Wörterbuch ist das ANW, welche literarischen Quellen hat es benutzt, wie viele lexikalisierte Wortgruppen hat es? Welche Artikel wurden vom Lexikografen X gemacht? Wie viele belgisch-niederländische (flämische) Wörter enthält es, welche Abbildungen sind im Wörterbuch aufgenommen? Das Suchgebiet des Computers ist hier nicht nur das Wörterbuch, sondern auch die Titelei und die verschiedenen Dokumentations- und Hilfetexte des Wörterbuchs und des ANW-Korpus.

4. Abschluss

Das ANW ist zurzeit (Ende 2007) mit den Vorbereitungen für die Publikation der Resultate aus der ersten Bearbeitungsphase im Internet beschäftigt. Mit Informatikern der Abteilung „Automatisierung“ des Instituts für niederländische Lexikologie wird an dem so genannten funktionellen Entwurf gearbeitet, der im April 2008 fertig sein soll. Dann folgen der technische Entwurf, der grafische Entwurf und der Bau eines Prototyps der Anwendung. Im November beginnt das Stadium der Realisierung, Implementation und der Tests. Anfang Mai 2009 soll das ANW im Internet zugänglich sein – und somit eine weitere Übereinstimmung mit *ellexiko* haben.

5. Literatur

- Moerdijk, Fons (2007a): Een nieuwe generatie elektronische woordenboeken. In: *Philologia Frisica Anno 2005. Lezingen fan it sechtjinde Frysk filologekongres 14, 15 en 16 desimber 2005*. Leeuwarden. S. 195-217.
- Moerdijk, Fons (2007b): Definities, frames en semagrammen. In: Moerdijk, Fons/Tempelaars, Rob/Van Santen, Ariane (Hg.): *Leven met woorden*. Leiden. S. 63-75.

6. Nachschlagewerke

- GWDS = Duden – Das große Wörterbuch der deutschen Sprache (2000). 10 Bde. auf CD-ROM. Mannheim.
- ioED = Oxford English Dictionary Online: <http://dictionary.oed.com/>
- iTLF = Le Trésor de la Langue Française informatisé: <http://atilf.atilf.fr/tlf.htm>
- iWNT = Woordenboek der Nederlandsche Taal op internet: <http://www.wnt.inl.nl>