Compiling medical dictionaries and spellcheckers for the Dutch language

Arnoud van den Eerenbeemt

Keywords: *medical lexicography*, *innovation on methodology*, *automation of lexicographical tasks*.

Abstract

This article outlines general aspects of medical monolingual lexicography in Dutch as based on the author's personal experience since 1995, converting the typesetting file of the Dutch *Pinkhof* monolingual medical dictionary into a dictionary database, editing the database since then, updating spelling and definitions and compiling specialised pocket dictionaries, spellcheckers and even medical dictates from the lexical content.

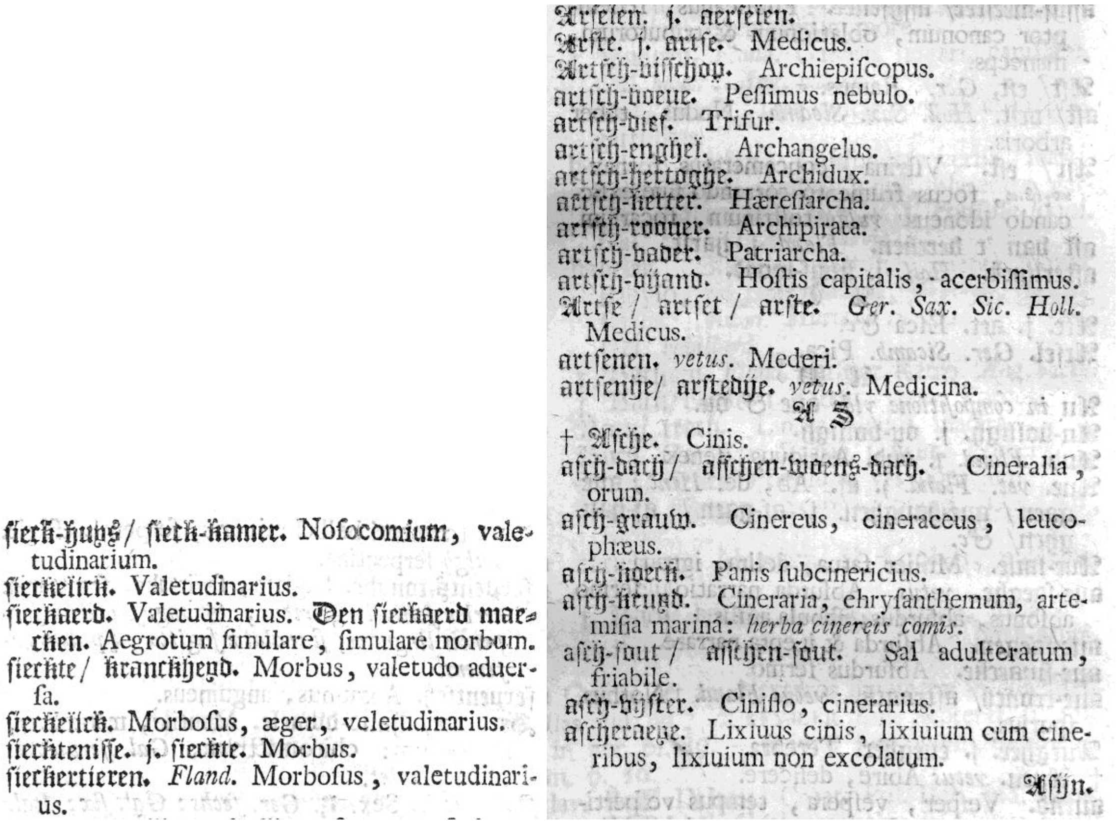
1. Medical Dutch

In Dutch healthcare, some 300,000 medical, paramedical and perimedical students and professionals communicate in 'medical Dutch'. Their idiom is a blend of Dutch and foreign medical official terms and jargon, to some extent borrowed from English, Latin, Greek, German and French. It disambiguates and standardises, speeds up communication, confirms competence and status and challenges and impresses with shibboleths such as: "Nurse, the skin disorder **rosacea** is to be pronounced *roSAcea*, not *rosaCEa*, please." Medical students are required to familiarise themselves with medical Latin replacing many common Dutch words: "Forget **sore** and **sores**, remember and pronounce correctly **ulcus** and **ulcera**. No more **suicide attempts**, but **tentamina suicidii**. Not **a heart disease**, but **cor vitium**." And so on.

The hybrid nature of medical idiom makes users prone to errors related to:

- *semantics:*
    - medicine is faced with an incessant stream of new terms and changing definitions;
    - meanings may differ depending on language and location ('false friends');
- *spelling:*
    - many terms differ only slightly in spelling, for example *two ureteral ducts* and *one urethral duct*, located at either side of the bladder;
    - many Latin terms are paired with Dutch bastard forms, for example *asthma* and *astma*; Latin is mastered by few medical professionals nowadays;
    - a term's spelling become less familiar once an acronym is replacing the full term;
    - long and complex terms feature multiple nouns and adjectives, interspersed with brackets, dashes, Greek letters and other symbols;
- *pronunciation*, with rules differing strongly for Dutch, (Greco-)Latin and English.

Therefore, mono- and multilingual medical dictionaries and spellcheckers have ample reasons to exist. Already in 1599, Kiliaan's first dictionary of the Dutch language featured strikingly many medical terms (figure 1).



ſieck-huys / ſieck-kamer. Noſocomium, valetudinarium.
ſieckelick. Valetudinarius.
ſieckaerd. Valetudinarius. Den ſieckaerd maercken. Aegrotum ſimulare, ſimulare morbum.
ſieckte / kranckheyd. Morbus, valetudo aduerſa.
ſieckelick. Morboſus, æger, veletudinarius.
ſiecktenisſe. j. ſieckte. Morbus.
ſieckertieren. *Fland.* Morboſus, valetudinarius.

Arſelen. j. aerſelen.
Arſte. j. artſe. Medicus.
Artſch-biſſchop. Archiepiſcopus.
artſch-boeue. Peſſimus nebulo.
artſch-dief. Trifur.
artſch-enghel. Archangelus.
artſch-hertoghe. Archidux.
artſch-ketter. Hæreſiarcha.
artſch-rooner. Archipirata.
artſch-vader. Patriarcha.
artſch-vijand. Hoſtis capitalis, acerbiſſimus.
Artſe / artſet / arſte. *Ger. Sax. Sic. Holl.* Medicus.
artſenen. *vetus.* Mederi.
artſenije/ arſtedije. *vetus.* Medicina.

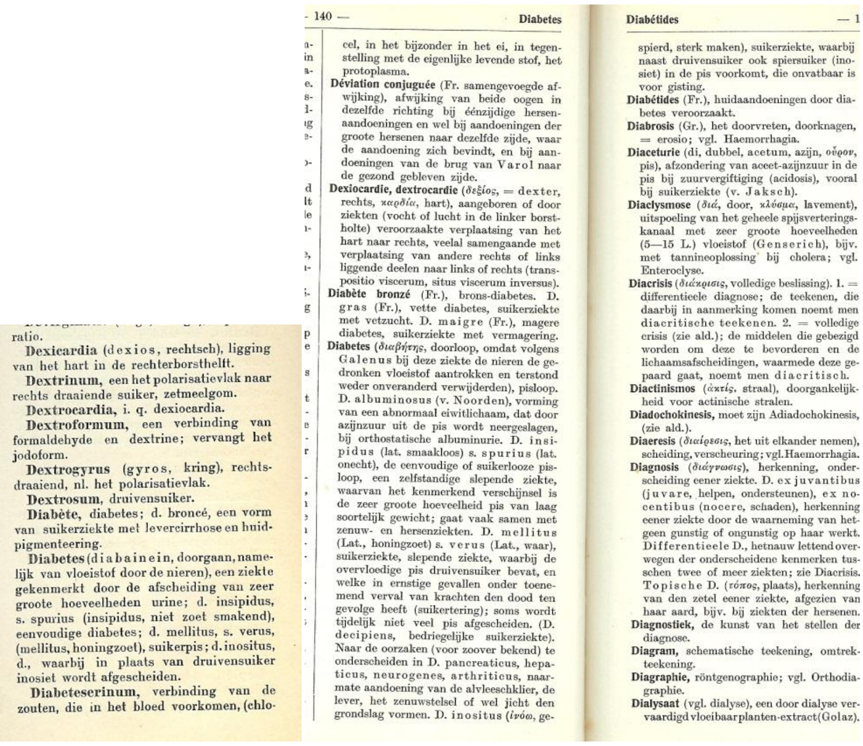
A S

† Aſche. Cinis.
aſch-dach / aſſchen-woens-dach. Cineralia, orum.
aſch-grauw. Cinereus, cineraceus, leucophæus.
aſch-koeck. Panis ſubcinericius.
aſch-kruyd. Cineraria, chryſanthemum, artemiſia marina: *herba cinereis comis.*
aſch-ſout / aſſchen-ſout. Sal adulteratum, friabile.
aſch-vijſter. Ciniflo, cinerarius.
aſcheraeghe. Lixiuus cinis, lixiuium cum cineribus, lixiuium non excolatum.

Aſijn.

Figure 1. Kiliaan, the first Dutch lexicographer, including quite a few medical terms in his *Etymologicum Teutonicae Linguae sive Dictionarium Teutonico-Latinum* (1599).

Some minor Dutch medical glossaries were published in the late 19th century, but only in 1923 physician Herman Pinkhof published his *Vertalend en verklarend woordenboek van uitheemsche geneeskundige termen* (‘Translating and explanatory dictionary of foreign medical terms’; from here on referred to as *Pinkhof*). This first comprehensive dictionary for medical students and professionals has since then been published by Bohn publishers, known since 1990 as Bohn Stafleu van Loghum and meanwhile part of Springer B+M.



ratio.

Dexicardia (dexios, rechtsch), ligging van het hart in de rechterborsthelft.

Dextrinum, een het polarisatievlak naar rechts draaiende suiker, zetmeelgom.

Dextrocardia, i. q. dexiocardia.

Dextroformum, een verbinding van formaldehyde en dextrine; vervangt het jodoform.

Dextrogyrus (gyros, kring), rechtsdraaiend, nl. het polarisatievlak.

Dextrosum, druivensuiker.

Diabète, diabetes; d. broncé, een vorm van suikerziekte met levercirrhose en huidpigmenteering.

Diabetes (diabainein, doorgaan, namelijk van vloeistof door de nieren), een ziekte gekenmerkt door de afscheiding van zeer groote hoeveelheden urine; d. insipidus, s. spurius (insipidus, niet zoet smakend), eenvoudige diabetes; d. mellitus, s. verus, (mellitus, honingzoet), suikerpis; d. inositus, d., waarbij in plaats van druivensuiker inosiet wordt afgescheiden.

Diabeteserinum, verbinding van de zouten, die in het bloed voorkomen, (chlo-

cel, in het bijzonder in het ei, in tegenstelling met de eigenlijke levende stof, het protoplasma.

Déviation conjuguée (Fr. samengevoegde afwijking), afwijking van beide oogen in dezelfde richting bij éénzijdige hersenaandoeningen en wel bij aandoeningen der groote hersenen naar dezelfde zijde, waar de aandoening zich bevindt, en bij aandoeningen van de brug van Varol naar de gezond gebleven zijde.

Dexiocardie, dextrocardie (δεξίος, = dexter, rechts, καρδία, hart), aangeboren of door ziekten (vocht of lucht in de linker borstholte) veroorzaakte verplaatsing van het hart naar rechts, veelal samengaande met verplaatsing van andere rechts of links liggende deelen naar links of rechts (transpositio viscerum, situs viscerum inversus).

Diabète bronzé (Fr.), brons-diabetes. D. gras (Fr.), vette diabetes, suikerziekte met vetzucht. D. maigre (Fr.), magere diabetes, suikerziekte met vermagering.

Diabetes (διαβήτης, doorloop, omdat volgens Galenus bij deze ziekte de nieren de gedronken vloeistof aantrokken en terstond weder onveranderd verwijderden), pisloop. D. albuminosus (v. Noorden), vorming van een abnormaal eiwitlichaam, dat door azijnzuur uit de pis wordt neergeslagen, bij orthostatische albuminurie. D. insipidus (lat. smaakloos) s. spurius (lat. onecht), de eenvoudige of suikerlooze pisloop, een zelfstandige slepende ziekte, waarvan het kenmerkend verschijnsel is de zeer groote hoeveelheid pis van laag soortelijk gewicht; gaat vaak samen met zenuw- en hersenziekten. D. mellitus (Lat., honingzoet) s. verus (Lat., waar), suikerziekte, slepende ziekte, waarbij de overvloedige pis druivensuiker bevat, en welke in ernstige gevallen onder toenemend verval van krachten den dood ten gevolge heeft (suikertering); soms wordt tijdelijk niet veel pis afgescheiden. (D. decipiens, bedriegelijke suikerziekte). Naar de oorzaken (voor zoover bekend) te onderscheiden in D. pancreaticus, hepaticus, neurogenes, arthriticus, naarmate aandoening van de alvleeschklier, de lever, het zenuwstelsel of wel jicht den grondslag vormen. D. inositus (ἰνόω, ge-

spierd, sterk maken), suikerziekte, waarbij naast druivensuiker ook spiersuiker (inosiet) in de pis voorkomt, die onvatbaar is voor gisting.

Diabétides (Fr.), huidaandoeningen door diabetes veroorzaakt.

Diabrosis (Gr.), het doorvreten, doorknagen, = erosio; vgl. Haemorrhagia.

Diaceturie (di, dubbel, acetum, azijn, οὖρον, pis), afzondering van aceet-azijnzuur in de pis bij zuurvergiftiging (acidosis), vooral bij suikerziekte (v. Jaksch).

Diaclysmose (διά, door, κλύσμα, lavement), uitspoeling van het geheele spijsverteringskanaal met zeer groote hoeveelheden (5—15 L.) vloeistof (Genserich), bijv. met tannineoplossing bij cholera; vgl. Enteroclyse.

Diacrisis (διάκρισις, volledige beslissing). 1. = differentieele diagnose; de teekenen, die daarbij in aanmerking komen noemt men diacritische teekenen. 2. = volledige crisis (zie ald.); de middelen die gebezigd worden om deze te bevorderen en de lichaamsafscheidingen, waarmede deze gepaard gaat, noemt men diacritisch.

Diactinismos (ἀκτίς, straal), doorgankelijkheid voor actinische stralen.

Diadochokinesis, moet zijn Adiadochokinesis, (zie ald.).

Diaeresis (διαίρεσις, het uit elkander nemen), scheiding, verscheuring; vgl. Haemorrhagia.

Diagnosis (διάγνωσις), herkenning, onderscheiding eener ziekte. D. ex juvantibus (juvare, helpen, ondersteunen), ex nocentibus (nocere, schaden), herkenning eener ziekte door de waarneming van hetgeen gunstig of ongunstig op haar werkt. Differentieele D., hetnauw lettend overwegen der onderscheidene kenmerken tusschen twee of meer ziekten; zie Diacrisis. Topische D. (τόπος, plaats), herkenning van den zetel eener ziekte, afgezien van haar aard, bijv. bij ziekten der hersenen.

Diagnostiek, de kunst van het stellen der diagnose.

Diagram, schematische teekening, omtrekteekening.

Diagraphie, röntgenographie; vgl. Orthodiagraphie.

Dialysaat (vgl. dialyse), een door dialyse vervaardigd vloeibaarplanten-extract(Golaz).

Figure 2. The term 'diabetes' described in Gabler, *Latijnsch geneeskundig woordenboek* (1918) and in the first edition of Pinkhof, *Vertalend en verklarend woordenboek van uitheemsche geneeskundige termen* (1923).

2. Medical lexicography

Relatively few lexicographers have specialised in medicine. To my disappointment, I have never met a medical lexicographer during four Euralex conferences attended. A large country as Germany boasts two comprehensive monolingual medical dictionaries only. Belgium doesn't even have a medical dictionary of its own and relies on *Pinkhof*.

A medical dictionary aimed for professional use will typically cover:

- *fundamental subjects*, such as anatomy, embryology, pathology, biochemistry, cell biology/histology, immunology, clinical genetics, pharmacology, physiology, statistics, epidemiology, medical law and medical informatics and E-healthcare;
- *clinical fields*, for example surgery, gynaecology, dermatology, nephrology, some 30 specialties altogether;
- *interdisciplinary fields*, such as paediatrics, geriatrics, travellers' medicine and social medicine;
- *paramedical and perimedical subjects*, such as dentistry, physiotherapy, nursing and midwifery.

Its users are expected to be familiar with terms from biology, (bio)chemistry, physics etc. used in definitions. Main aspects for defining a disease are anatomy (where is it located?), etiology (what is its cause?), diagnostics (how was it found?), pathology (what goes wrong in the body?), treatment and prognosis. A Dutch medical term's meaning may depend on the medical specialty involved.

3. The *Pinkhof* dictionary

The current, 12th edition of *Pinkhof* defines in encyclopaedic definitions approximately 53,000 terms, with 6,000 newly added terms. 3500 entries feature a note on usage and some 2000 terms have been marked *confusable*. Thirty articles discuss general topics on medical grammar, semantics and pragmatics in Dutch versus Latin and English, such as “when does a bug name end on -i and when on -ii?”, “why is ‘processi spinosi’ not the plural form of ‘processus spinosus’?” and “when am I supposed to spell ‘asthma’ or ‘astma’ in Dutch?”.

3.1. *Database-aided Pinkhof lexicography*

Pinkhof’s ninth edition, published in 1992, had been prepared in a simple ad-hoc MS-DOS database with limited features. The editorial board recognised the need for a genuine database serviced by a linguistic assistant. In 1995 I was hired as the first full-time editor and was teamed up with dermatologist Van Everdingen, since 1983 editor-in-chief on a part-time free-lance basis. Since then I have been working on both monolingual and bilingual medical dictionaries, spellcheckers and even medical dictates and quizzes. To my relief, medical specialists consulted assure me that by now my terminographical experience is largely compensating for my lack of a medical education. Nevertheless, respecting the boundaries of my knowledge and competence is essential in my profession.

Figure 3. Pinkhof authors Van Everdingen and Van den Eerenbeemt (note that the stitching on his shirt reads *lexicon ordbok*).

3.2. *Pinkhof's macro- and microstructure*

Pinkhof's macrostructure consists of an elaborate introductory part, with a recommendation by a key opinion leader, a preface, listing of referents and a detailed instruction on how to use the book, followed by the 53,000 articles in the main part. A term may be found as both a main entry and a sub entry, but the definition will be found under the sub entry only. There are no illustrations.

As for data displayed in print, the microstructure is: headword, headword language, form variant(s), form variant language, optional spelling form(s), abbreviated/full form, field label(s), definition(s), usage notes, synonym(s), synonym language, *see [also]* reference(s); biographical details, etymology, grammaticalia, pronunciation, confusable term(s).

Additionally, the main administrative, non-displayed fields are: sort term, sort phrase, homograph field, editorial remarks, source, specialty code 1-6, creation/editing date, name of editor and print-yes/no.

Creating a record for a new term often involves adding additional records for its Dutch/Latin form variant(s), its acronym, its synonym(s), all these referring to the mother record with *see*. Though monolingual, *Pinkhof* will occasionally mention English equivalents common in Dutch healthcare as a synonym. Where appropriate I will add grammaticalia, etymology and usage comments, including sample phrases and semantic and grammatical pitfalls.

In May 2012, my pronunciation of 1160 medical tongue twisters was recorded for inclusion in the app and PC editions. A video fragment showing me in the recording studio pronouncing Nordic medical eponyms was posted on Google+ as a contest question and made Dutch followers send in various attempts to determine the spoken medical entries.

3.3. *Role of nomenclatures and spelling regulations*

The spelling, meaning and usage of medical terms are rather changeable as research worldwide leads to an incessant expansion and revision of terms. A medical nomenclature committee may decide to revise all names of microorganisms or anatomical entities every few years or so.

- In 1998 FICAT anatomists revised 9,000 terms on the human body and replaced the official nomenclature *Nomina Anatomica* with *Terminologia Anatomica*. Some 1000 terms had been added or deleted and even changes in Latin anatomical spelling were made. In 2012, senior doctors are still not familiar with the 'TA'.
- In English-speaking countries, anatomical Latin terms are increasingly being replaced with their English equivalents.
- Even microorganisms get a new name now and then, for example ***Pneumocystis carinii*** being renamed in ***Pneumocystis jiroveci***. Yet the acronym for the pneumonia caused by it is still **PCP**, not **PJP**.
- Increasingly, eponyms are replaced with terms providing insight in localisation and pathology, such as **spondylitis ankylopoetica** or **ankylosing spondylitis** instead of **Bechterew's disease**.

3.4. *Medical spelling in Pinkhof*

Spelling rules for Dutch were revised drastically in 1995 and slightly once more in 2005, requiring me to revise thousands of Dutch eponym compounds, among other things.

In 2011 the Dutch WHO-FIC committee translating the International classification of diseases (ICD-10) invited me to develop a new Dutch medical spelling in order to increase spelling consistency. The translators had observed that applying spelling rules may lead to erratic spelling forms in complex medical terms. A custom spelling existed already for Dutch biologists, allowing them to write *de Kraai in de Eik* (*the Crow in the Oak*) rather than *de kraai en de eik*. This inspired me to work out, in cooperation with various spelling authorities, additional rules offering medical professionals a choice between the official spelling in *maak 2 12 afleidingen-ecg's bij dit brugadasyndroom* and 'medical-idiom spelling' in *maak 2 12-afleidingen-ECG's bij dit Brugada-syndroom* (translated loosely: 'make two 12-electrode ECGs for this Brugada syndrome'). The latter form avoids misunderstandings and corresponds better with spelling in English. The new spelling rules are optional and apply to some 0.5 per cent of all *Pinkhof* entries only. The additional spelling form is clearly marked and explained in detail in *Pinkhof*.

4. Methodology

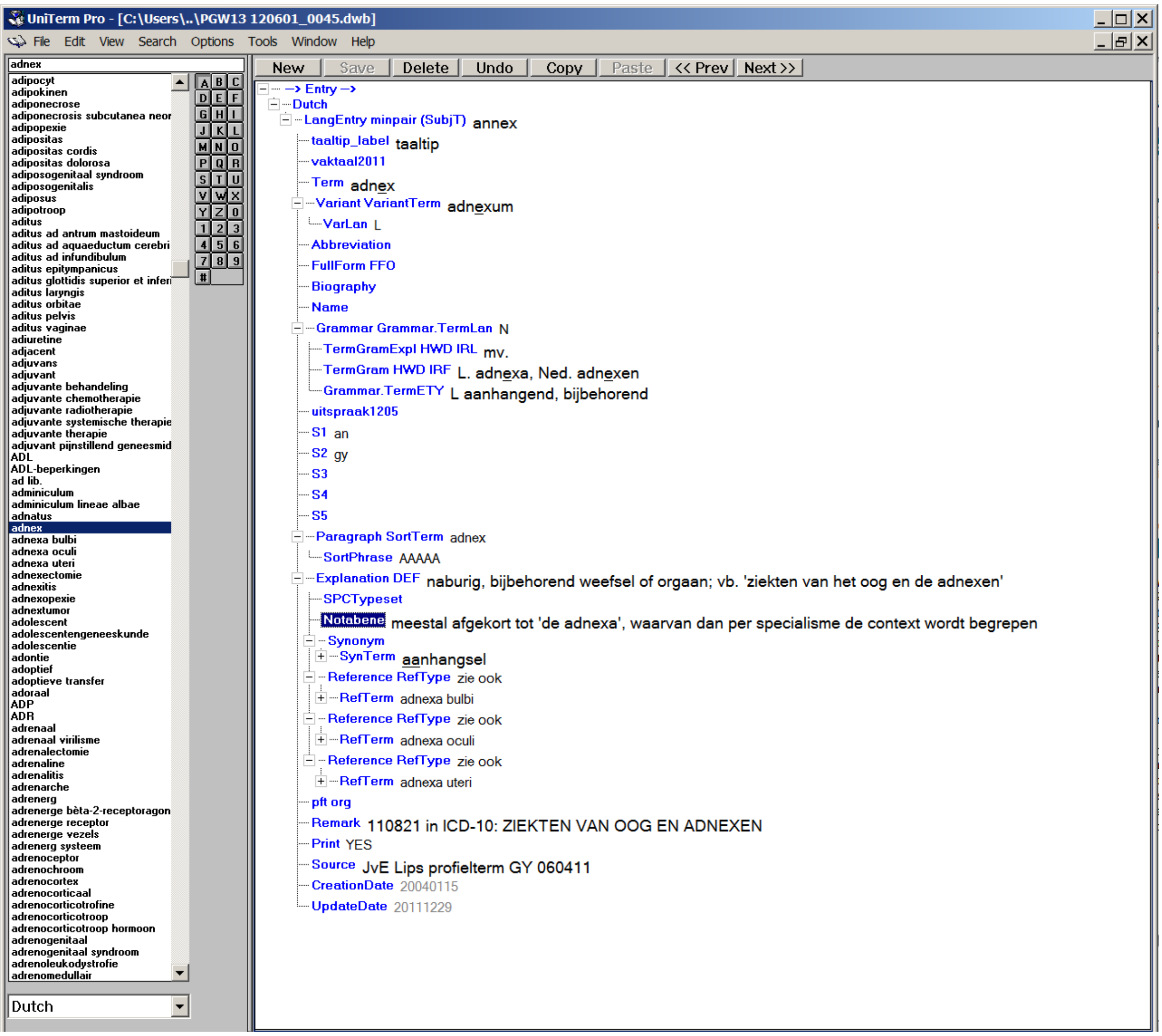
4.1. *Converting from typesetting files to a database*

After joining the editorial board in 1995 and preparing myself by reading's Sidney Landau's *Dictionaries: the Art and Craft of Lexicography*, my prime task was to convert the typesetting file into a structured and tagged text file, to find suitable dictionary software, assign lexemes and metadata to fields and import the result into a database. Once the 42,000 records had been imported into CX-Master (developed by Henning Madsen, Compulexis Ltd.), 9000 entries were marked *obsolete* and removed. Thus we disposed of for instance '**absinthism**: addiction to absinth', an addictive liquor banned in our country nearly a century ago, as well as of spelling characteristics abolished in 1948 (*stroo*, *visch*) and quite a number of other obsolete words and entries.

For a major update we then invited some thirty medical students. Those passing a writing test were instructed to scan index entries of 90 selected medical text books used by Dutch medical universities for important terms missing in the dictionary. The resulting 9,000 terms were described with the description in the text book serving as the basis for a dictionary definition. For this purpose they received an elaborate manual on medical terminography listing both the main types of medical definitions and linguistic pitfalls to be avoided. About forty medical specialists and I were to review all concept definitions. This allowed me to familiarise myself thoroughly with both medical terminology and terminography.

In that period, Dutch lexicologist Raffaela Vlot wrote her master thesis on *Pinkhof* under the supervision of lexicologist prof. Willy Martin. She joined me in my office for three months, examined the dictionary's microstructure and discussed her lexicological observations with prof. Martin and me. Our joint findings have contributed to the overall quality of the dictionary.

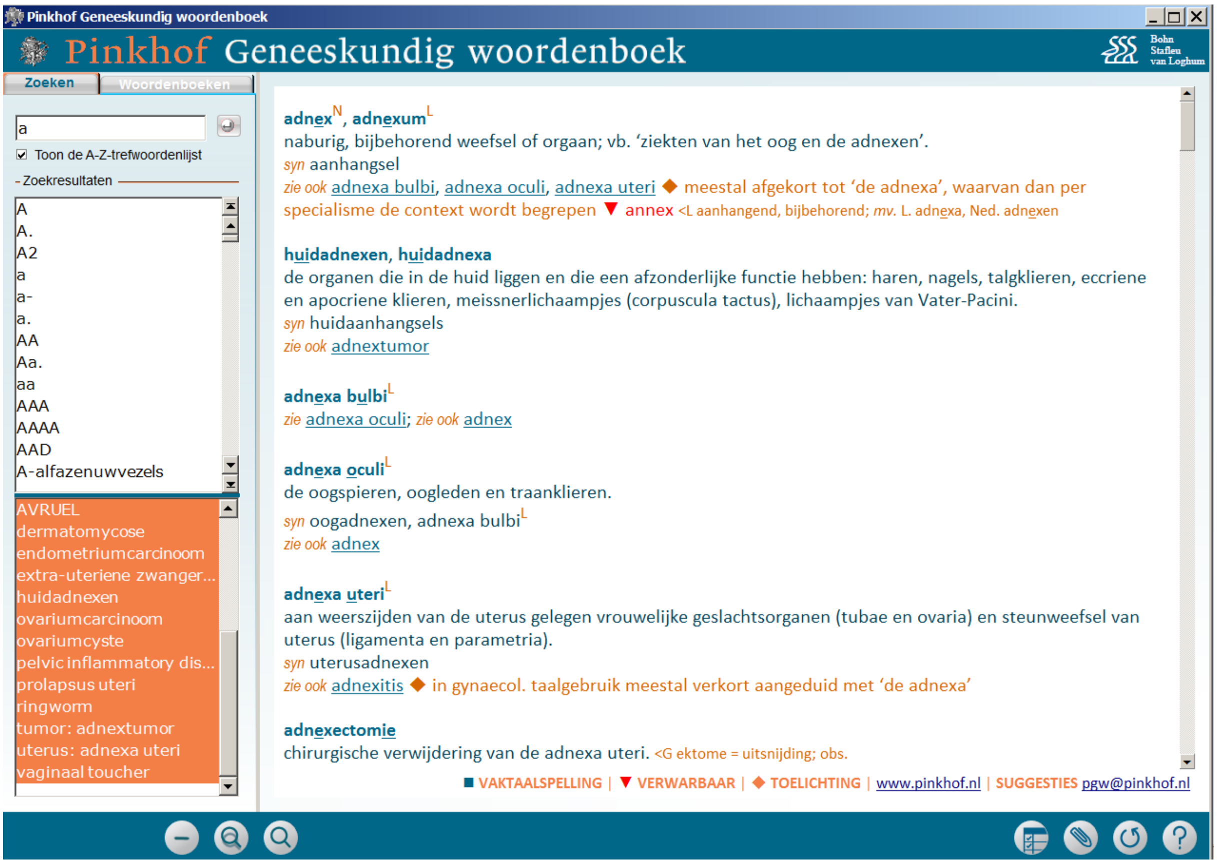
Once the need of working in Windows was felt in 2004, I changed to Uniterm Pro (Acolada GmbH).


UniTerm Pro - [C:\Users\..\PGW13 120601_0045.dwb]
File
Edit
View
Search
Options
Tools
Window
Help
New
Save
Delete
Undo
Copy
Paste
<< Prev
Next >>
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z 0 1 2 3 4 5 6 7 8 9 #

```
adnex
adipocyt
adipokinen
adiponecrose
adiponecrosis subcutanea neor
adipopexie
adipositas
adipositas cordis
adipositas dolorosa
adiposogenitaal syndroom
adiposogenitalis
adiposus
adipotroop
aditus
aditus ad antrum mastoideum
aditus ad aquaeductum cerebri
aditus ad infundibulum
aditus epitympanicus
aditus glottidis superior et inferi
aditus laryngis
aditus orbitae
aditus pelvis
aditus vaginae
adiuretine
adjacent
adjuvans
adjuvant
adjuvante behandeling
adjuvante chemotherapie
adjuvante radiotherapie
adjuvante systemische therapie
adjuvante therapie
adjuvant pijnstillend geneesmid
ADL
ADL-beperkingen
ad lib.
adminiculum
adminiculum lineae albae
adnatus
adnex
adnexa bulbi
adnexa oculi
adnexa uteri
adnexectomie
adnexitis
adnexopexie
adnextumor
adolescent
adolescentengeneeskunde
adolescentie
adontie
adoptief
adoptieve transfer
adoraal
ADP
ADR
adrenaal
adrenaal virilisme
adrenalectomie
adrenaline
adrenalitis
adrenarche
adrenerg
adrenerge bèta-2-receptoragon
adrenerge receptor
adrenerge vezels
adrenerg systeem
adrenoceptor
adrenochroom
adrenocortex
adrenocorticaal
adrenocorticotrofine
adrenocorticotroop
adrenocorticotroop hormoon
adrenogenitaal
adrenogenitaal syndroom
adrenoleukodystrofie
adrenomedullair
Dutch
```

```
--> Entry -->
  Dutch
    LangEntry minpair (SubjT) annex
      taaltip_label taaltip
      vaktaal2011
      Term adnex
      Variant VariantTerm adnexum
        VarLan L
      Abbreviation
      FullForm FFO
      Biography
      Name
      Grammar Grammar.TermLan N
        TermGramExpl HWD IRL mv.
        TermGram HWD IRF L. adnexa, Ned. adnexen
        Grammar.TermETY L aanhangend, bijbehorend
      uitspraak1205
      S1 an
      S2 gy
      S3
      S4
      S5
      Paragraph SortTerm adnex
        SortPhrase AAAAA
      Explanation DEF naburig, bijbehorend weefsel of orgaan; vb. 'ziekten van het oog en de adnexen'
        SPCTypeset
        Notabene meestal afgekort tot 'de adnexa', waarvan dan per specialisme de context wordt begrepen
        Synonym
          SynTerm aanhangsel
        Reference RefType zie ook
          RefTerm adnexa bulbi
        Reference RefType zie ook
          RefTerm adnexa oculi
        Reference RefType zie ook
          RefTerm adnexa uteri
      pft org
      Remark 110821 in ICD-10: ZIEKTEN VAN OOG EN ADNEXEN
      Print YES
      Source JvE Lips profielterm GY 060411
      CreationDate 20040115
      UpdateDate 20111229
```

Figure 4. The UniTerm Pro editing screen showing the Pinkhof adnex entry.


Pinkhof Geneeskundig woordenboek
Pinkhof Geneeskundig woordenboek
Bohn Stafleu van Loghum
Zoeken
Woordenboeken
a
☑ Toon de A-Z-trefwoordenlijst
Zoekresultaten

A
A.
A2
a
a-
a.
AA
Aa.
aa
AAA
AAAA
AAD
A-alfazenuwvezels

AVRUEL
dermatomycose
endometriumcarcinoom
extra-uteriene zwanger...
huidadnexen
ovariumcarcinoom
ovariumcyste
pelvic inflammatory dis...
prolapsus uteri
ringworm
tumor: adnextumor
uterus: adnexa uteri
vaginaal toucher

adnex^{N}, **adnexum**^{L}
naburig, bijbehorend weefsel of orgaan; vb. 'ziekten van het oog en de adnexen'.
syn aanhangsel
zie ook adnexa bulbi, adnexa oculi, adnexa uteri ◆ meestal afgekort tot 'de adnexa', waarvan dan per specialisme de context wordt begrepen ▼ annex <L aanhangend, bijbehorend; *mv.* L. adnexa, Ned. adnexen

huidadnexen, huidadnexa
de organen die in de huid liggen en die een afzonderlijke functie hebben: haren, nagels, talgklieren, eccriene en apocriene klieren, meissnerlichaampjes (corpuscula tactus), lichaampjes van Vater-Pacini.
syn huidaanhangsels
zie ook adnextumor

adnexa bulbi^{L}
zie adnexa oculi; *zie ook* adnex

adnexa oculi^{L}
de oogspieren, oogleden en traanklieren.
syn oogadnexen, adnexa bulbi^{L}
zie ook adnex

adnexa uteri^{L}
aan weerszijden van de uterus gelegen vrouwelijke geslachtsorganen (tubae en ovaria) en steunweefsel van uterus (ligamenta en parametria).
syn uterusadnexen
zie ook adnexitis ◆ in gynaecol. taalgebruik meestal verkort aangeduid met 'de adnexa'

adnexectomie
chirurgische verwijdering van de adnexa uteri. <G ektome = uitsnijding; obs.

■ VAKTAALSPELLING | ▼ VERWARBAAR | ◆ TOELICHTING | www.pinkhof.nl | SUGGESTIES pgw@pinkhof.nl

Figure 5. The *Pinkhof* PC edition showing the **adnex** entry.

4.2. *Enhancing editing methods*

As Dutch is spoken by some 20 million people only, the potential sales of a medical dictionary will be limited and so is the publisher's budget. I am therefore required to develop methods in order to minimize farming out proofing tasks. I regularly assign Acolada engineers to write custom macros for complex batch tasks, such as creating records from synonyms, creating references, finding cyclic references, assignment of variable values, copying stressed syllables from 15,000 headword fields to 20,000 fields still lacking these, and so on.

In addition to Uniterm Pro I rely on:

- the text editor TextPad for powerful sorting, selecting and searching/replacing with regular expressions;
- from the WordSmith suite the tools WordList (for sorting terms backwards) and Minimal Pairs, developed by Mike Scott at my request for finding confusing lookalikes;
- speech recognition software for automating tedious tasks;
- advanced time-saving editing features in Microsoft Word known to too few users;
- ExamDiff Pro for detailed comparisons of listings.

As for speech recognition, in 2009 I translated *Fachwortschatz Medizin*, a bilingual medical learner's dictionary. With voice-driven macro commands I instructed my PC to switch between five applications displayed permanently, create and fill out Uniterm records, look up terms in various sources and make Google perform some 600,000 complex conditional searches. I thus avoided keying in the equivalent of 6000 pages of text. Moreover, whilst going through the roughly 42,000 German headwords and 100,000 English translations I was able to export some 2,000 newly found terms into *Pinkhof*.

4.3. *Validating content*

4.3.1 *Validating spelling.* My main resources for validating a term's spelling in Dutch are:

- the leading medical weekly journal *Netherlands Journal of medicine* (NTvG), consulted ad hoc on-line and additionally its paper edition read in full;
- the Dutch edition of the *International Classification of Diseases* (ICD-10);
- spelling rules prescribed by the Dutch Language Union, described for professionals in full detail in the *Technical [Spelling] Manual*;
- the Dutch online pharmacopeia *Farmacotherapeutisch Kompas*;
- the *Terminologia Anatomica* (Thieme Verlag) listing all 9000 terms in anatomical Latin;
- treatment guidelines of the Dutch association of practitioners (Nederlands Huisartsengenootschap) – and so on.

Some reference files have been pinned in my MS Word configuration for instant access. *Pinkhof* has been matched against all 1400 topics in the internal *NTvG* guide on spelling and style. *Pinkhof*'s middle-sized sole competitor is ignored as its spelling choices are inconsistent. I do not follow any nomenclature blindly as it may contain spelling inconsistencies. For instance, after screening some 5,000 names in the official virus nomenclature with WordList I was able to convince its editors of several typos. The Dutch ICD-10, still in progress, contains spelling errors.

The WHO-FIC, the Dutch spelling committee and the *NTvG* consult me on medical spelling issues regularly.

4.3.2 *Validating definitions.* Main resources for validating definitions are:

- *Netherlands Journal of medicine* (NTvG);
- medical specialists consulted;
- Dutch medical textbooks;
- foreign medical dictionaries and websites such as *Pschyrembel*, *Stedman* and *Dorland*.

Co-editor Van Everdingen is involved in the development of medical guidelines in the Netherlands and has thus access to a network of medical specialists. They are invited to revise limited numbers of terms and definitions. We cannot expect these professionals, who contribute on an honorary basis, to be competent in spelling rules, nomenclature or database techniques. I therefore facilitate their work by limiting the number of entries (e.g. exclusively on juvenile diabetes, hand surgery or epilepsy) and supplying adequately formatted Word files. Sometimes I may include a listing of terms sorted on the final syllables, 'backwards'. In medicine, this will typically group terms that are related semantically and give more oversight.

5. *Pinkhof* thriving on-line

As of 2012, *Pinkhof* boasts five editions: the book, a PC edition (UniLex for Windows), the browser edition (UniLex IDS) on www.pinkhof.eu, an iOS app and an Android app. Digital editions boast many advantages. The book edition has become less attractive for both publisher and users. As exceeding its current 1520 pages would have doubled the binding costs, I was forced to remove functional blank lines from the book's design in order to reduce the number of pages. With digital editions medical students need no longer complain about heavy books. They may misspell Latin in a search string and yet have fuzzy search technology find the correct term. The PC edition features full-text searching (autumn 2012 also in UniLex IDS). Narrowing a search to eponyms or diagnostic terms, for example, was supported in the PDA edition (2006), but will feature again in the PC version 1.2. Online distribution allows for regular updates at a low cost. Supporting colours were not cost-effective in the latest book edition, but are available at no cost in a digital edition. The fully functional trial version may be downloaded at no cost from www.pinkhof.nl and be tried for 7 days. We intend to distribute the first *Pinkhof* edition (1923) as a searchable give-away PDF for promotional purposes.

One drawback of digital dictionaries, however, is seldom mentioned. I find serendipity a likeable aspect of browsing dictionaries in folio and therefore hope to include in some future update a start-up option displaying some random entry.

6. Repurposing medical lexical content in spin-offs

6.1. *Pocket dictionaries*

Pharmaceutical companies are in constant need of gifts relevant to the medical profession in order to promote drugs. The budget for this purpose is not meagre. I have so far compiled a score of specialised pocket editions, some of which have been translated into English, French and Spanish. My sales colleagues typically invite me to compile a dictionary on a specific indication, for example rheumatoid arthritis. I examine the characteristics of the drug and its class, relevant patho(physio)logy, diagnostics and therapy, and will then select relevant records on joint anatomy, pharmacology, histo- and immunopathology. This yields typically some 1000 out of 54,000 entries. I make sure that hyperonyms of selected subentries are included and that all *see (also)* references remain valid.

Each database record has one up to six medical specialty codes, allowing for a quick selection. However, the coding is not unambiguous and requires multiple search rounds, as a term on rheumatoid arthritis may have been assigned *1 rheumatology 2 immunology 3 surgery* or equally well *1 immunology 2 surgery 3 rheumatology*.

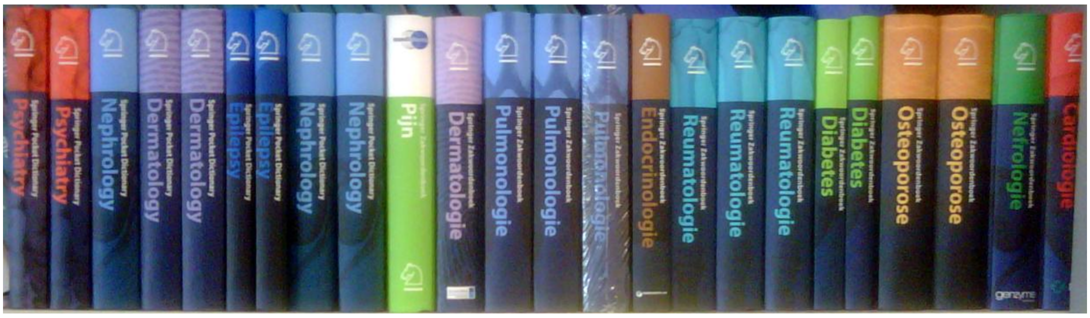


Figure 6. Pocket dictionaries extracted from the main *Pinkhof* dictionary.

6.2. *Spellcheckers*

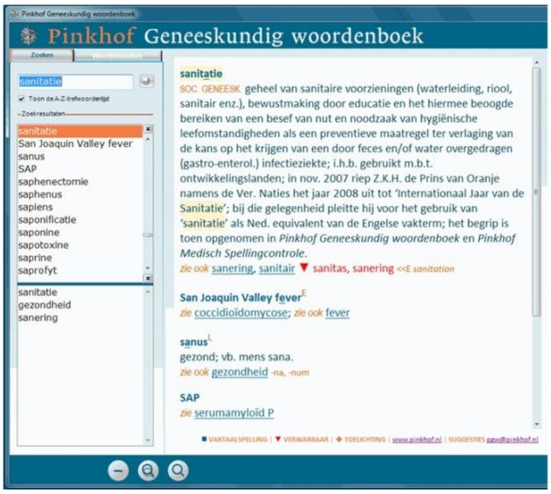
After the release of the 10th edition in 1998, extracted from a database for the first time, I started experimenting with WordSmith analysing content. Having now access to all medical words and morphemes I decided to develop a medical spellchecker, not yet available for the Dutch language. For a linguist this was definitely a challenging and fun assignment, as compiling a spellchecker requires a profound knowledge of ‘medicolinguistic morphology’.

WordSmith and TextPad allowed me to extract sort in various ways some 80.000 unique medical words that are not included in Word’s Dutch LEX file list. I enriched the content morphologically by adding inflected forms needed in Dutch and Latin for plural/singular, definite/indefinite etc. For compiling the MS Office add-in I consulted language engineers specialised in building proofing tool add-ins. They instructed me to add metadata on word class, compositionality and hyphenation. Eventually some 150,000 medical words were made available for *Pinkhof Medische spellingcontrole* version 1.0 on CD-ROM. At present version 4.1, a download version, comprises some 180,000 words. Its trial version with some 25 per cent of the content is available as a free download.

The spellchecker’s user interface, accessible through a custom button in the Microsoft Word ribbon, offers a word wheel enabling the user to look up medical terms with wild cards as well as the option to consult a comprehensive help text on spelling and grammar rules for medical Dutch.

Meanwhile I have compiled various specialised spellcheckers at the request of sales colleagues servicing pharmaceutical companies. As the words and morphemes included have no specialty code I cannot quickly compile a subset. Each ‘specialised’ spellchecker therefore simply contains the full spellchecker content plus a substantial number of specialised words, extracted for the occasion from pharmacological documentation supplied.

In 2008 the Dutch crown prince suggested including **sanitatie** in spellcheckers; his wish was my demand.


Pinkhof Geneeskundig woordenboek
Pinkhof Geneeskundig woordenboek
Zoeken
sanitatie
☑ Toon de A-Z-trefwoordenlijst
Zoekresultaten

sanitatie
San Joaquin Valley fever
sanus
SAP
saphenectomie
saphenus
sapiens
saponificatie
saponine
sapotoxine
saprine
saprofyt

sanitatie
gezondheid
sanering

sanitatie
SOC. GENEESK. geheel van sanitaire voorzieningen (waterleiding, riool, sanitair enz.), bewustmaking door educatie en het hiermee beoogde bereiken van een besef van nut en noodzaak van hygiënische leefomstandigheden als een preventieve maatregel ter verlaging van de kans op het krijgen van een door feces en/of water overgedragen (gastro-enterol.) infectieziekte; i.h.b. gebruikt m.b.t. ontwikkelingslanden; in nov. 2007 riep Z.K.H. de Prins van Oranje namens de Ver. Naties het jaar 2008 uit tot ‘Internationaal Jaar van de Sanitatie’; bij die gelegenheid pleitte hij voor het gebruik van ‘sanitatie’ als Ned. equivalent van de Engelse vakterm; het begrip is toen opgenomen in *Pinkhof Geneeskundig woordenboek* en *Pinkhof Medisch Spellingcontrole*.
zie ook sanering, sanitair ▼ sanitas, sanering <<E *sanitation*

San Joaquin Valley fever[E]
zie coccidioïdomycose; *zie ook* fever

sanus[L]
gezond; vb. mens sana.
zie ook gezondheid -na, -num

SAP
zie serumamyloïd P

■ VAKTAALSPELLING | ▼ VERWARBAAR | ◆ TOELICHTING | www.pinkhof.nl | SUGGESTIES pgw@pinkhof.nl

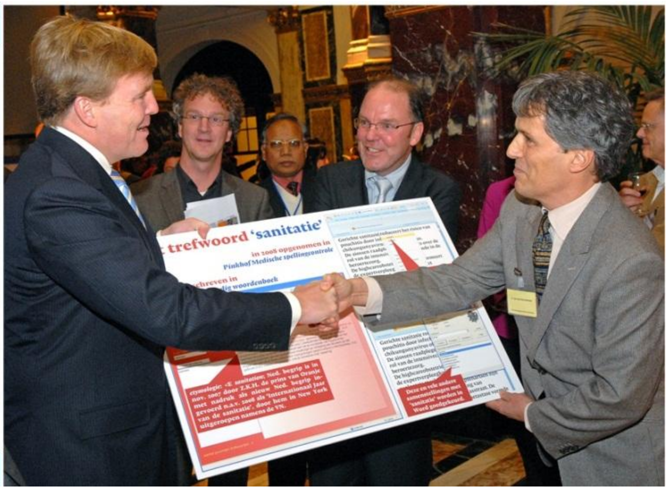


Figure 7. Introducing in 2008 my medical spellchecker to the Dutch crown prince, HRH Willem-Alexander van Oranje.

6.3. *Even… dictates and quizzes*

Some medical professionals are painfully aware of their poor knowledge of spelling rules and will, nevertheless, not hesitate to participate for fun in a medical dictate contest. Pharmaceutical companies organise these contests as a promotional event and have me write imaginary specialist's letters reporting on examinations of patients. Hosting a medical dictate contest at a conference for pulmonologists in San Francisco in 2009 was definitely an exciting event for a Dutch lexicographer accustomed to a secluded office environment.

From 2012 onward, I will be writing and hosting a quiz on medical language for students from all Dutch medical universities participating in the annual Rosalind Franklin Contest. My multiple-choice questions will confront both candidates and the participating audience alike with orthographic and semantic pitfalls in medicine.

Figure 8. The author whilst video-hosting a medical dictate, notably in front of office wallpaper with enlarged text taken from the Pinkhof medical dictionary.

7. Various aspects

7.1. *'Lexicographia pro juventute'*

In 2009 I was required to change my methodology drastically when Van Dale invited co-author Van Everdingen and me to write a medical dictionary for teenagers and their parents or guardians: *Van Dale Junior Dokterswoordenboek*. I am normally expected to abridge and upscale definitions, replacing lay terms with medical vernacular. Now I had to do quite the opposite: selecting and describing in plain Dutch 2500 diseases relevant to this target group, along with diagnostic examinations, treatments, biochemistry, drugs etc. Having inherited the proper glossopoetic genes from my mother, an author of children's books in the 60s, I embraced this task with pleasure and to the satisfaction of my co-author, the reviewers, the publisher and many readers.

7.2. *Promoting Pinkhof: Twitter and Google+*

Pinkhof is aimed primarily at medical students. As the *Pinkhof* author I need to stay in touch with their terminological needs and linguistic skills and idiosyncrasies. In addition to events such as the annual Rosalind Franklin Contest I contribute to promotional actions on a leading website for medical students. In 2011 we posted a contest online inviting students to look up answers in *Pinkhof* and rewarding them with free software licenses. As a part of www.pinkhof.nl an action game was developed requiring students to ‘perform surgery’ on a patient whilst answering multiple-choice questions.

As of January 2011 I have been using Twitter, extending this recently to Google+, for posting medicolinguistic observations. In a year’s period I have produced approximately 1,400 tweets on medical terminology and acquired some 300 avid followers. Every day I spend some fifteen minutes interacting with my target groups and enjoying their feedback and retweets. For this same purpose of community building, readers are invited to send suggestions to pgw@pinkhof.bsl on each book page, each PC edition screen and the Info screen in the apps.

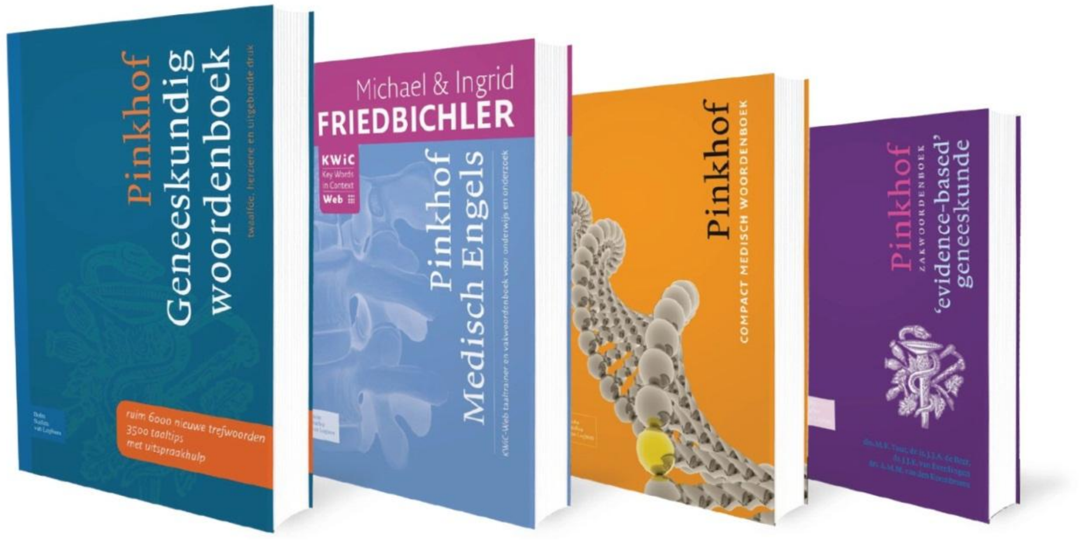
7.3. *Educating a museum’s audience*

At present I am preparing lexical exhibits on medical Dutch for the Leyden-based Boerhaave National Museum on the history of Medicine. To my pleasure, this respected institute has recognised the need for a display exhibiting linguistic aspects of Dutch medicine, among other aspects of dictionaries, the morphology of medical terms, nosology and classification systems.

7.4. *Medical ghost terms*

I always include in *Pinkhof* a few nonsensical medical terms and definitions as ghost terms in order to prevent plagiarism by announcing this in the colophon. In 1998, students sold for some € 150 a diskette with a MS Word file containing some 20,000 *Pinkhof* headwords (containing 3% of typos!) as a ‘medical spellchecker’. Unfortunately, they adhered too accurately to *Pinkhof*’s sorting, so it was a piece of cake to sue their company and have it closed down.

I trust that this last section will not discourage fellow lexicographers now becoming interested from venturing into medical lexicography.



References

Eerenbeemt, A.M.M. van den 2009. *Pinkhof Medische spellingcontrole*. (PC software, downloadable at www.pinkhof.nl). Houten: Bohn Stafleu van Loghum.

Everdingen, J.J.E and A. van den Eerenbeemt 2011. *Pinkhof Geneeskundig woordenboek*. (12th edition folio 2011; PC/app/browser edition 2012.) Houten: Bohn Stafleu van Loghum.

Everdingen, J.J.E and A. van den Eerenbeemt 2010. *Van Dale Junior Dokterswoordenboek*. Utrecht: Van Dale Uitgevers.

Friedbichler, M. en I. 2010. *Pinkhof Medisch Engels*. (Book and PC edition.) Translated from German into Dutch by A.M.M. van den Eerenbeemt. Houten: Bohn Stafleu van Loghum.

Gabler. *Latijnsch geneeskundig woordenboek.* (Bibliographical data incomplete, publication year unknown).

Kiliaan. *Etymologicum Teutonicae Linguae sive Dictionarium Teutonico-Latinum* (1599).

Pinkhof, H. 1923. *Vertalend en verklarend woordenboek van uitheemsche geneeskundige termen*. (2nd through 12th edition: 1932-2012.) Haarlem: Bohn.

www.pinkhof.nl

Social media (in Dutch):

Twitter	@dokterstaal
Google+	‘Pinkhof – medisch taalvaardig’