

# LINGUISTICS WITH CLARIN SEARCH ILLUSTRATION 1

Jan Odijk

LOT Winterschool

Amsterdam, 2015-01-12

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Example Problem (based on [Odijk 2011](#))
- Glimpse of
  - Searching in PoS-tagged Corpus
  - Searching for grammatical relations
  - Searching for Constructions
  - Searching for synonyms/ hyponyms
  - Analyzing/Visualising Word occurrence patterns in CHILDES

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

[MORE](#)

	A	P	V
	Zij is daar __ blij mee	Zij is daar __ mee in haar nopjes	Zij verheugde zich daar __ over
Zeer	OK	OK	OK
Erg	OK	OK	OK
Heel	OK	*	*

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Differences
  - not due to semantics
  - purely syntactic
  - does not follow from a general principle,
  - so it must be ‘learned’ by a child acquiring Dutch as a first language

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Research Questions
  - How can such facts be acquired (L1 acquisition)?
  - How can child learn that *zeer* and *heel* can modify A, V, and P?
    - Is there enough evidence for this to the child?
  - How can a child `learn' that *heel* **cannot** modify Ps or Vs-> there is no evidence for this (no negative evidence)
    - Is there a relation between time of acquisition and modification potential?
    - Role of indirect negative evidence?
- (and much more can be said about this)

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- How to approach this problem
  - Study literature, study grammars, form and test hypotheses, **look for relevant data sets, create new datasets, enrich data with annotations, search in and through datasets, analyze data and visualize analysis results**, design and carry out experiments, design and do simulations, ....
  - Focus here: searching relevant data easily in large resources using (components of ) the CLARIN infrastructure

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Google is [no good](#) for this!
  - Because you need (inter alia) grammatical information
  - Because (as any decent word) the relevant words are highly ambiguous (syntax and semantics):
    - Erg (4x)= noun(de) 'erg'; noun(het)'evil', adj+adv 'unpleasant', adv 'very'
    - Zeer (3x)= noun 'pain'; adj 'painful'; adv 'very'
    - Heel (4x) = adj 'whole'; adj 'big'; verbform 'heal'; adv 'very'

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Are the basic facts correct?
- Search with [OpenSONAR](#)
  - Search in PoS-tagged corpus SONAR-500
  - reduces problem with ambiguities
  - Sneak preview
- [Demo](#)



# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Conclusions after analysis
  - *Heel* does occur with certain adverbially used PPs
    - *Heel in het begin, heel af en toe, heel in het bijzonder, heel in het kort, heel op het laatst, heel in de verte, heel uit de verte, heel in het algemeen,*
    - *Dat ligt hem heel na aan het hart*
  - *Heel* does occur with predicative PPs (but I find them ill-formed)
    - *buiten zijn verwachting, in de mode, in de vakantiestemming, in het zwart, in orde*
  - Maybe *heel* is used as *geheel* by some people

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- PoS code annotation
  - is (just) OK for adjacent words (but quite some noise)
  - Is useless for more distant grammatically related words
- Desired: Search for words that have a grammatical relation (dependency relations)
- [LASSY Woordrelaties Interface](#)
  - LASSY Small: 65 k sentences (1 m words)
  - LASSY-LARGE/wiki: 8.6 m sentences (125 m words)
- [Demo](#)

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Conclusions
  - *Heel*
    - There are examples where *heel* modifies a `verb`
    - But `verb` is actually a deverbal (participle) adjective
    - in '*heel open staan voor*' *heel* is incorrectly analyzed as modifying the verb
  - *Zeer*:
    - most examples of deverbal adjectives
    - But also some real verbs
  - → confirms initial assumptions about the facts

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Searching for Constructions
  - [GrETEL](#)
  - Example-based treebank query system
    - LASSY-Small, Corpus Gesproken Nederlands (CGN)
    - Recently extended to SONAR (500 m tokens)

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- [Cornetto data](#) and [Interface](#) to Cornetto
- Lexico-semantic database based on Dutch WordNet and ReferentieBestand Nederlands
- Created in STEVIN programme
- User-friendly interface made in CLARIN-NL
- [Example](#) to search for (near-)synonyms of *zeer*, *erg*, *heel*.

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- What is the modification potential of near-synonyms of *zeer*, *heel*, *erg*?
  - *allemachtig-adv-2 beestachtig-adv-2 bijzonder-a-4 bliksems-adv-2 bloedig-adv-2 bovenmate-adv-1 buitengewoon-adv-2 buitenmate-adv-1 buitensporig-adv-2 crimineel-a-4 deerlijk-adv-2 deksels-adv-2 donders-adv-2 drommels-adv-2 eindeloos-a-3 enorm-adv-2 erbarmelijk-adv-2 fantastisch-adv-6 formidabel-adv-2 geweldig-adv-4 goddeloos-adv-2 godsjammerlijk-adv-2 grenzeloos-adv-2 grotelijks-adv-1 heel-adv-5 ijselijk-adv-2 izig-a-4 intens-adv-2 krankzinnig-adv-3 machtig-adv-4 mirakels-adv-1 monsterachtig-adv-2 moorddadig-adv-4 oneindig-adv-2 onnoemelijk-adv-2 ontegelijk-adv-2 ontstellend-adv-2 ontzaglijk-adv-2 ontzettend-adv-3 onuitsprekelijk-adv-2 onvoorstelbaar-adv-2 onwezenlijk-adv-2 onwijs-adv-4 overweldigend-adv-2 peilloos-adv-2 reusachtig-adv-3 reuze-adv-2 schrikkelijk-adv-2 sterk-adv-7 uiterst-adv-4 verdomd-adv-2 verdraaid-a-4 verduiveld-adv-2 verduveld-adv-2 verrekt-adv-3 verrot-adv-3 verschrikkelijk-adv-3 vervloekt-adv-2 vreselijk-adv-5 waanzinnig-adv-2 zeer-adv-3 zeldzaam-adv-2 zwaar-adv-10*
- Many of these appear atypical for young children and are probably learned late
- Is there a correlation between this and their modification potential?

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- [COAVA](#) application CHILDES browser
  - Application built for research into the relation between language acquisition and lexical dialectical variation
  - Cognition, Acquisition and Variation tool
  - [Demo](#) of the COAVA CHILDES browser analyzing and visualising children's speech
  - (for child-directed speech see [here](#))

# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

	found	mod A	mod V	mod N	mod P	other	Un-clear
<b>zeer</b>	52	<b>1</b>	0	0	<b>0</b>	51	
<b>heel</b>	800	<b>744</b>	4	7	<b>0</b>	2	43
<b>erg</b>	54	25	1	1	<b>0</b>	26	1

First relevant occurrence	heel	erg	zeer
Day(Yr;Mo)	705 (1;11)	1048 (2;10)	1711 (4;8)



# CLARIN INFRASTRUCTURE

## TOOLS: ILLUSTRATION

- Summary: CLARIN-NL tools
  - Enable search for grammatical and semantic properties
  - In small (1M) to large (500M) annotated corpora
  - And in rich lexical databases
  - With easy to use interfaces
  - Provide new data gathering opportunities
    - that mostly did not exist for Dutch until recently
    - were available for specialists only until one year ago

Thanks for your attention!

DO NOT ENTER HERE

NL

A large, faint, light gray network diagram in the background. It consists of several circular nodes connected by thin lines. One node at the top right is highlighted with a white border and contains the letters "NL". The diagram is composed of several overlapping circular arcs and straight lines connecting the nodes.

# GOOGLE V. DESIRED



Property	Google	What you want
String search	yes	yes
Relation between strings	nearness	Grammatical relations, PoS codes
Search for function words	No / unreliable	Yes
Search for morpho-syntactic and syntactic properties	no	Yes
Construction search	no	Yes
Dutch only	unreliable	Yes
Size	huge	Huge (but so far there is only small (1m) or large (700m))

# Virtual Language Observatory

Explore the world of language resources and technology from different perspectives



[VLO Home](#) >> [Faceted Browser Resources](#)

## COLLECTION

- [Meertens collection: Liederenbank](#) (243129)
- [DK-CLARIN Repository](#) (46658)
- [Nederlands Instituut voor Beeld en Geluid Academia collectie](#) (46156)
- [childes](#) (28594)
- [DoBeS archive](#) (21162)
- [Language and Cognition](#) (20792)
- [Institut für Deutsche Sprache, CLARIN-D Zentrum, Mannheim](#) (14320)
- [talkbank](#) (14243)
- [Acquisition](#) (13142)
- [MPI CGN](#) (12767)
- [more...](#)

## CONTINENT

- [Europe](#) (64567)
- [North-America](#) (21669)
- [Asia](#) (17084)
- [South-America](#) (8231)
- [Oceania](#) (5073)
- [Africa](#) (5053)
- [Australia](#) (2287)
- [Middle-America](#) (2175)

## LANGUAGE

- [Dutch](#) (140814)
- [Nederlands](#) (109557)
- [English](#) (61570)
- [German](#) (41116)
- [Danish](#) (26553)
- [French](#) (11641)
- [Spanish; Castilian](#) (11258)
- [Deutsch](#) (7149)
- [Japanese](#) (7009)
- [Turkish](#) (6241)
- [more...](#)

## GENRE

- [discourse](#) (83462)
- [kinderlied](#) (18594)
- [primary\\_text](#) (7523)
- [liefdeslied](#) (5812)
- [bespiegeland lied](#) (4722)
- [newspaper\\_text](#) (4410)
- [stimuli](#) (3965)
- [story\\_telling](#) (3912)

Showing 1 to 10 of 545258 << < 1 2 3 4 5 6 7 8 9 10 >> >>

name	description
<a href="#">"DE STIEFMOEDER" die haar 18 jarigen STIEFZOON 10 jaar heeft opgesloten op de 3e verdieping Over-Amstelstraat 10 te AMSTERDAM</a>	
<a href="#">"DE STIEFMOEDER" die haar 18 jarigen STIEFZOON 10 jaar heeft opgesloten op de 3e verdieping Over-Amstelstraat 10 te AMSTERDAM</a>	
<a href="#">"T IS FEEST."</a>	
<a href="#">"T IS FEEST."</a>	
<a href="#">"ACH, MIJN LIEVE AUGUSTIJN."</a>	
<a href="#">"ACH, MIJN LIEVE AUGUSTIJN."</a>	
<a href="#">"ACH, MIJN LIEVE AUGUSTIJN."</a>	
<a href="#">"Barberstue i ""Punch"" "</a>	Scene fra barberstue. "Punch". Negativ og aftryk gave fra magister Hans Lassen. Foto som forarbejde...
<a href="#">"Barberstue i ""Punch"" "</a>	Scene fra barberstue. "Punch". Negativ og aftryk gave fra magister Hans Lassen. Foto som forarbejde...
<a href="#">"Barberstue i ""Punch"" "</a>	Barberstue "Punch" 4. 7. 1878. Negativ og aftryk gave fra magister Hans Lassen. Foto

Quick navigation to: [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [L](#) [M](#) [N](#) [O](#) [P](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#)

## A

[AILLA](#) (920)  
[ANDES](#) (233)  
[AVATech webservices](#) (5)  
[Acquisition](#) (13142)  
[Adaptive Listening Group](#) (1)  
[African Sign Languages](#) (432)  
[Arnhem Land Sign Languages](#) (1)  
[Arthurian Fiction project](#) (348)  
[Audio Archive of Linguistic Fieldwork](#) (199)  
[Austrian Sign Language](#) (111)

## B

[BAS SIGNUM](#) (28)  
[Bavarian Archive for Speech Signals \(BAS\)](#) (10602)  
[Boiste](#) (1)

## C

[CHILDES Data repository](#) (314)  
[CLARIN LRT inventory](#) (933)  
[CLARIN NL](#) (3375)  
[CLARIN-NL web services](#) (3)  
[CLLE-ERSS Université de Toulouse-Le Mirail](#) (7)  
[CMDI Providers](#) (235)  
[CORP-ORAL](#) (47)  
[Comparative Corpus of Spoken Portuguese](#) (1)  
[Comparative cognitive anthropology](#) (610)  
[Comprehension](#) (168)  
[Corpus NGT](#) (564)  
[childes](#) (28594)

## M

[MPI CGN](#) (12767)  
[MPI IMDI Archive](#) (2)  
[MPI für Bildungsforschung](#) (2919)  
[MUCH](#) (1)  
[Magoria Books' Carib and Romani Archive](#) (2)  
[Meertens Collection: Diversity in Dutch DP Design \(DiDDD\)](#) (334)  
[Meertens Collection: Dynamische Fonologische en Morfologische Atlas van de Nederlandse Dialecten \(GTRP\)](#) (614)  
[Meertens Collection: Dynamische Syntactische Atlas van de Nederlandse Dialecten \(DvnaSAND\)](#) (268)  
[Meertens Collection: GBA-derived](#) (1)  
[Meertens Collection: GBA-derived - sub COROP area](#) (40)  
[Meertens Collection: GBA-derived - sub dialectarea](#) (24)  
[Meertens Collection: GBA-derived - sub municipality](#) (443)  
[Meertens Collection: GBA-derived - sub province](#) (12)  
[Meertens Collection: Heiloo toponiemen](#) (399)  
[Meertens Collection: Soundbites](#) (1983)  
[Meertens collection: CRM](#) (96)  
[Meertens collection: Etstoel](#) (999)  
[Meertens collection: Liederenbank](#) (243129)  
[Meta-Share Repository at ILC](#) (89)  
[Multilingualism](#) (2361)  
[Multimodal Learning and teaching Corpora Exchange](#) (49)

## N

[NEHOL](#) (44)  
[Namescape: mapping the onymic landscape](#) (6)  
[Nederlands Instituut voor Beeld en Geluid Academia collectie](#) (46156)  
[Neurobiology of language](#) (275)  
[Nijmegen corpora of casual speech](#) (319)

# Virtual Language Observatory



Explore the world of language resources and technology from different perspectives



[VLO Home](#) >> [Faceted Browser Resources](#)

**COLLECTION :** [all](#) > CHILDES Data repository

**LANGUAGE**

- [English](#) (121)
- [Spanish; Castilian](#) (36)
- [French](#) (25)
- [Dutch](#) (19)
- [German](#) (19)
- [Japanese](#) (18)
- [Italian](#) (12)
- [Chinese](#) (11)
- [Portuguese](#) (10)
- [Hebrew](#) (8)
- [more...](#)

**CONTINENT**

**GENRE**

- [primary\\_text](#) (314)

**COUNTRY**

**SUBJECT**

- [child language development](#) (312)
- [child language development, language disorders](#) (2)

**ORGANISATION**

**RESOURCE TYPE**

Showing 1 to 10 of 314 << < 1 2 3 4 5 6 7 8 9 10 > >>

name	description
<a href="#">Afrikaans vanDulm Corpus</a>	longitudinal study of two Afrikaans children, L2 data also available
<a href="#">Basque Luque Corpus</a>	Three visits each to 38 Basque children aged 2 to 4
<a href="#">Basque Soto Valle Corpus</a>	Longitudinal lagged study of 8 Basque children
<a href="#">Bilingual Navracsics Corpus</a>	Persian, English, Hungarian trilingual children
<a href="#">Bilingual AarsenBos Corpus</a>	Arabic-Dutch bilinguals, Turkish-Dutch bilinguals, and monolingual control groups for Arabic,...
<a href="#">Bilingual CUHK Student Corpus</a>	conversations of children of various ages learning English, cannot be used for publications
<a href="#">Bilingual DeHouwer Corpus</a>	Longitudinal study of an English-Dutch bilingual child over an 8-month period
<a href="#">Bilingual Deuchar Corpus</a>	Longitudinal study of a Spanish-English bilingual child in England
<a href="#">Bilingual FerFuLice Corpus</a>	study of twins in Spain learning Spanish and English
<a href="#">Bilingual Genesee Corpus</a>	French-English bilingual children in Montreal

# Virtual Language Observatory

Explore the world of language resources and technology from different perspectives



[VLO Home](#) >> Faceted Browser Resources

 search

**COLLECTION** : [all](#) > CHILDES Data repository

**CONTINENT**

**COUNTRY**

**ORGANISATION**

[TalkBank](#) (19)

**DATAPROVIDER**

[OLAC Metadata Providers](#) (19)

**TAG**

**LANGUAGE** : [all](#) > Dutch

**GENRE**

[primary text](#) (19)

**SUBJECT**

[child language development](#) (19)

**RESOURCETYPE**

[text](#) (19)

[unknown type](#) (19)

**NATIONALPROJECT**

Showing 1 to 10 of 19

<< < 1 2 > >>

name	description
<a href="#">Bilingual AarsenBos Corpus</a>	Arabic-Dutch bilinguals, Turkish-Dutch bilinguals and monolingual control groups for Arabic,...
<a href="#">Bilingual DeHouwer Corpus</a>	Longitudinal study of an English-Dutch bilingual child over an 8-month period
<a href="#">Bilingual van Oosten Corpus</a>	Dutch-Italian bilinguals and monolingual comparison children, picture descriptions
<a href="#">Clinical Dutch BolKuiken Down Corpus</a>	Dutch Down children and adolescents
<a href="#">Clinical Dutch BolKuiken HI Corpus</a>	Dutch Hearing-Impaired children
<a href="#">Clinical Dutch BolKuiken Normal Corpus</a>	Dutch Normal children
<a href="#">Clinical Dutch BolKuiken SLI Corpus</a>	Dutch SLI children
<a href="#">Clinical Dutch BolPol PDD-NOS Corpus</a>	Dutch children with PDD-NOS pervasive developmental disorder, not otherwise specified
<a href="#">Dutch DeHouwer Corpus</a>	this study focused on dialect features unique to the Antwerp area
<a href="#">Dutch Gillis Corpus</a>	longitudinal corpus of bi-weekly recording sessions of a boy learning Dutch



# Virtual Language Observatory

Explore the world of language resources and technology from different perspectives



[back to results](#) | no original context available | [found an error?](#)

< Record 10 out of 19 >

Field	Value
<b>name</b>	Dutch Gillis Corpus
<b>description</b>	longitudinal corpus of bi-weekly recording sessions of a boy learning Dutch
<b>id</b>	oai:chilides.talkbank.org:Germanic-Dutch-Gillis
<b>collection</b>	CHILDES Data repository
<b>dataProvider</b>	OLAC Metadata Providers
<b>genre</b>	primary_text
<b>languages</b>	<a href="#">Dutch</a>
<b>metadataSource</b>	<a href="http://catalog.clarin.eu/oai-harvester/olac-and-dc-providers/harvested/results/cmdi/CHILDES_Data_repository/oai_chilides_talkbank_org_Germanic_Dutch_Gillis.xml">http://catalog.clarin.eu/oai-harvester/olac-and-dc-providers/harvested/results/cmdi/CHILDES_Data_repository/oai_chilides_talkbank_org_Germanic_Dutch_Gillis.xml</a>
<b>organisation</b>	TalkBank
<b>subject</b>	child language development

Resources:

 <http://chilides.talkbank.org/data-xml/Germanic/Dutch/Gillis.zip>

 <http://chilides.talkbank.org/data/Germanic/Dutch/Gillis.zip>

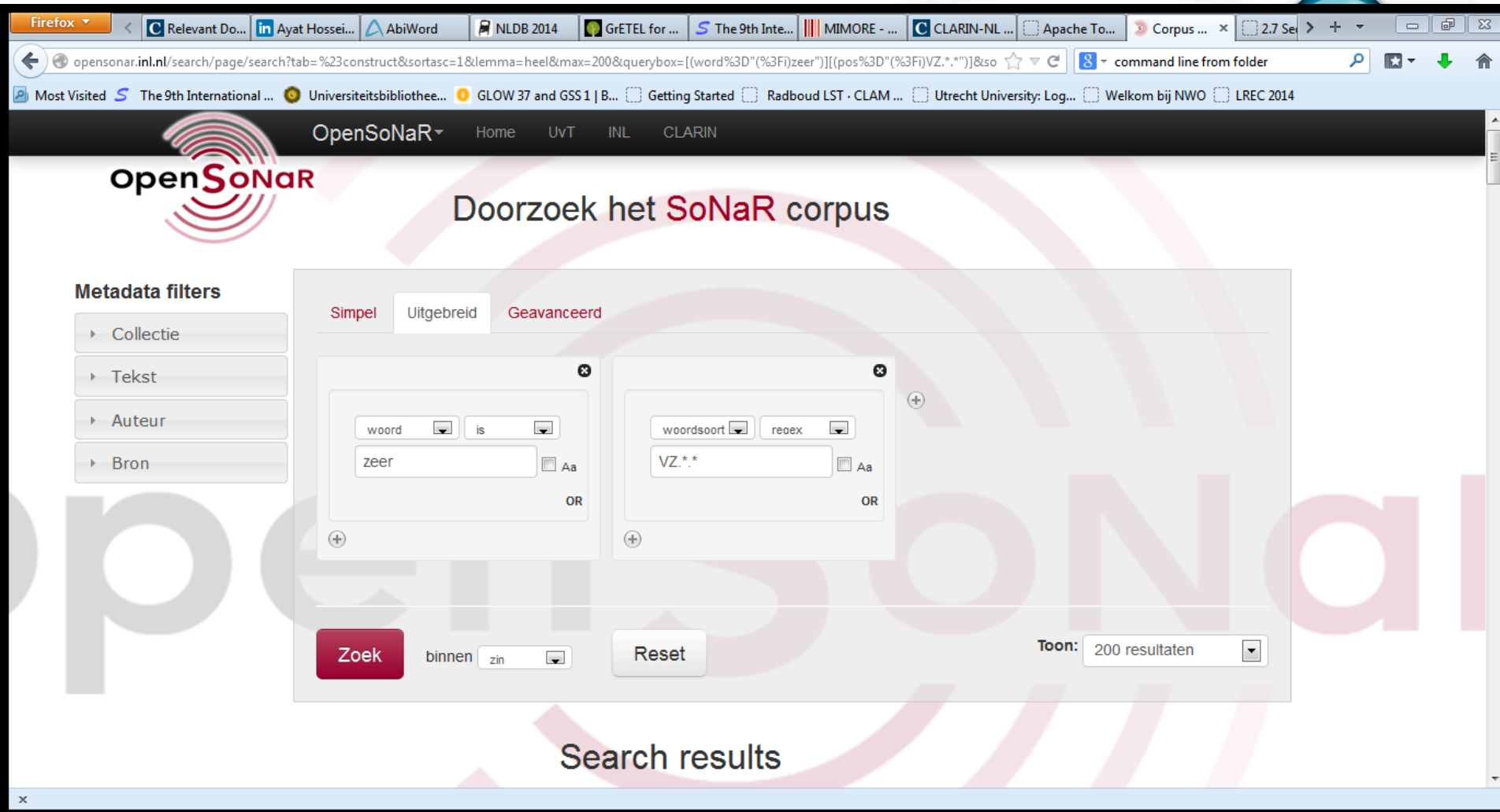
- RETURN Page



# OPENSONAR



- Start Page



Firefox

opensonar.in.nl/search/page/search?tab=%23construct&sortasc=1&lemma=heel&max=200&querybox=[(word%3D(%3F)zeer)][(pos%3D(%3F)VZ.\*\*)]&so

OpenSoNaR Home UvT INL CLARIN

## openSoNaR

### Doorzoek het SoNaR corpus

**Metadata filters**

- Collectie
- Tekst
- Auteur
- Bron

**Simpel** **Uitgebreid** Geavanceerd

woord is zeer Aa OR

woordsoort reoex VZ.\*\* Aa OR

Zoek binnen zin Reset

Toon: 200 resultaten

### Search results

Firefox | Relevant Do... | Ayat Hossei... | AbiWord | NLDB 2014 | GRETTEL for ... | The 9th Inte... | MIMORE - ... | CLARIN-NL ... | Apache To... | Corpus ... x | 2.7 Se | + | - | x

opensonar.inl.nl/search/page/search?tab=%23construct&sortasc=1&lemma=heel&max=200&querybox=[(word%3D"%3F)zeer"][(pos%3D"%3F)VZ.\*")]&so ☆ | command line from folder | ☆ | ↓ | 🏠

Most Visited | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS 1 | B... | Getting Started | Radboud LST - CLAM ... | Utrecht University: Log... | Welkom bij NWO | LREC 2014

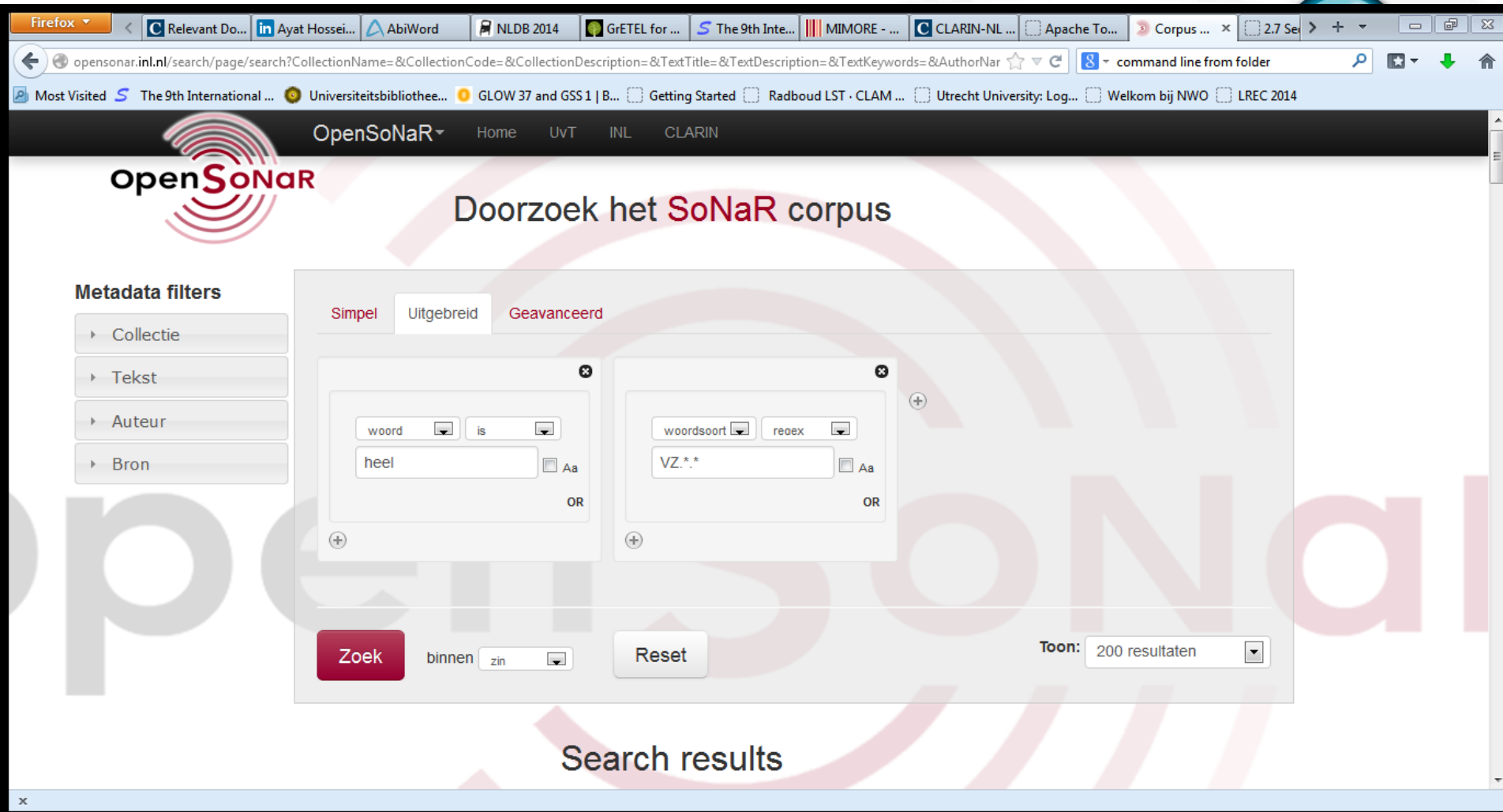
OpenSoNaR | Home | UvT | INL | CLARIN

**openSoNaR** Hits grouped Documents grouped

Total hits: 4227  
Total pages: 22

Prev 1 2 3 4 5 6 7 8 9 10 11 12 13 14 ... Next Toggle titles

Left context ▾ Hit text Right context ▾	Lemma	Part of speech
... ook buiten het Nederlandstalig gebied <b>zeer in</b> trek . Ogeschijnlijk was het ...	zeer in	BW() VZ(init)
... macht , alzo dat wij <b>zeer in</b> twijfel waren , ook van ...	zeer in	BW() VZ(init)
... als mijn moeder waren te <b>zeer in</b> goede doen om te vertrekken ...	zeer in	BW() VZ(init)
... gebeurtenissen bij de andere volken <b>zeer in</b> aanzien was gestegen , kostbare ...	zeer in	BW() VZ(init)
... dat we elkaar al te <b>zeer in</b> de weg zaten . Geleidelijk ...	zeer in	BW() VZ(init)
... had , maar hij was <b>zeer in</b> beslag genomen door de film ...	zeer in	BW() VZ(init)
... eenlettergrepige woorden ; ze leek <b>zeer in</b> beslag genomen door het gesprek ...	zeer in	BW() VZ(init)
... vreemde ; mijn nieren verlangen <b>zeer in</b> mijn schoot . 28 Voorwaar ...	zeer in	BW() VZ(init)
... , en Hij hem niet <b>zeer in</b> overvloed doorkend heeft ; 16 ...	zeer in	BW() VZ(init)
... , draag ik een diep <b>zeer in</b> mijn hart , dat er ...	zeer in	BW() VZ(init)
... en zij ontzeten zich bovenmate <b>zeer in</b> zichzelf , en waren verwonderd ...	zeer in	BW() VZ(init)
... kende , niet al te <b>zeer in</b> de wind geslagen te hebben ...	zeer in	BW() VZ(init)
... die op dit ogenblik weer <b>zeer in</b> het gedrang is gekomen , ...	zeer in	BW() VZ(init)
... deze tekst hem al te <b>zeer in</b> de richting gaat van het ...	zeer in	BW() VZ(init)
... autodidact , een kwalificatie die <b>zeer in</b> de mode was in de ...	zeer in	BW() VZ(init)
... moederschap : ik ben te <b>zeer in</b> mezelf gekeerd , gefocust op ...	zeer in	BW() VZ(init)



The screenshot shows the OpenSoNaR search interface in a Firefox browser. The browser's address bar contains the URL `opensonar.in.nl/search/page/search?CollectionName=&CollectionCode=&CollectionDescription=&TextTitle=&TextDescription=&TextKeywords=&AuthorNar`. The page title is "Doorzoek het SoNaR corpus".

**OpenSoNaR** Home UvT INL CLARIN

## Doorzoek het SoNaR corpus

**Metadata filters**

- Collectie
- Tekst
- Auteur
- Bron

**Search filters:**

- Simpel** (selected) | Uitgebreid | Geavanceerd
- Field: woord, Value: heel, Case: Aa, Operator: OR
- Field: is, Value: is, Case: Aa, Operator: OR
- Field: woordsoort, Value: VZ.\*\* (with asterisks), Case: Aa, Operator: OR
- Field: reoex, Value: reoex, Case: Aa, Operator: OR

**Search options:**

- Zoek (button)
- binnen: zin (dropdown)
- Reset (button)
- Toon: 200 resultaten (dropdown)

**Search results**

Firefox | Relevant Do... | Ayat Hossei... | AbiWord | NLDB 2014 | GrETEL for ... | The 9th Inte... | MIMORE - ... | CLARIN-NL ... | Apache To... | Corpus ... x | 2.7 Se... | + | - | x

opensonar.inl.nl/search/page/search?CollectionName=&CollectionCode=&CollectionDescription=&TextTitle=&TextDescription=&TextKeywords=&AuthorNar ☆ | command line from folder

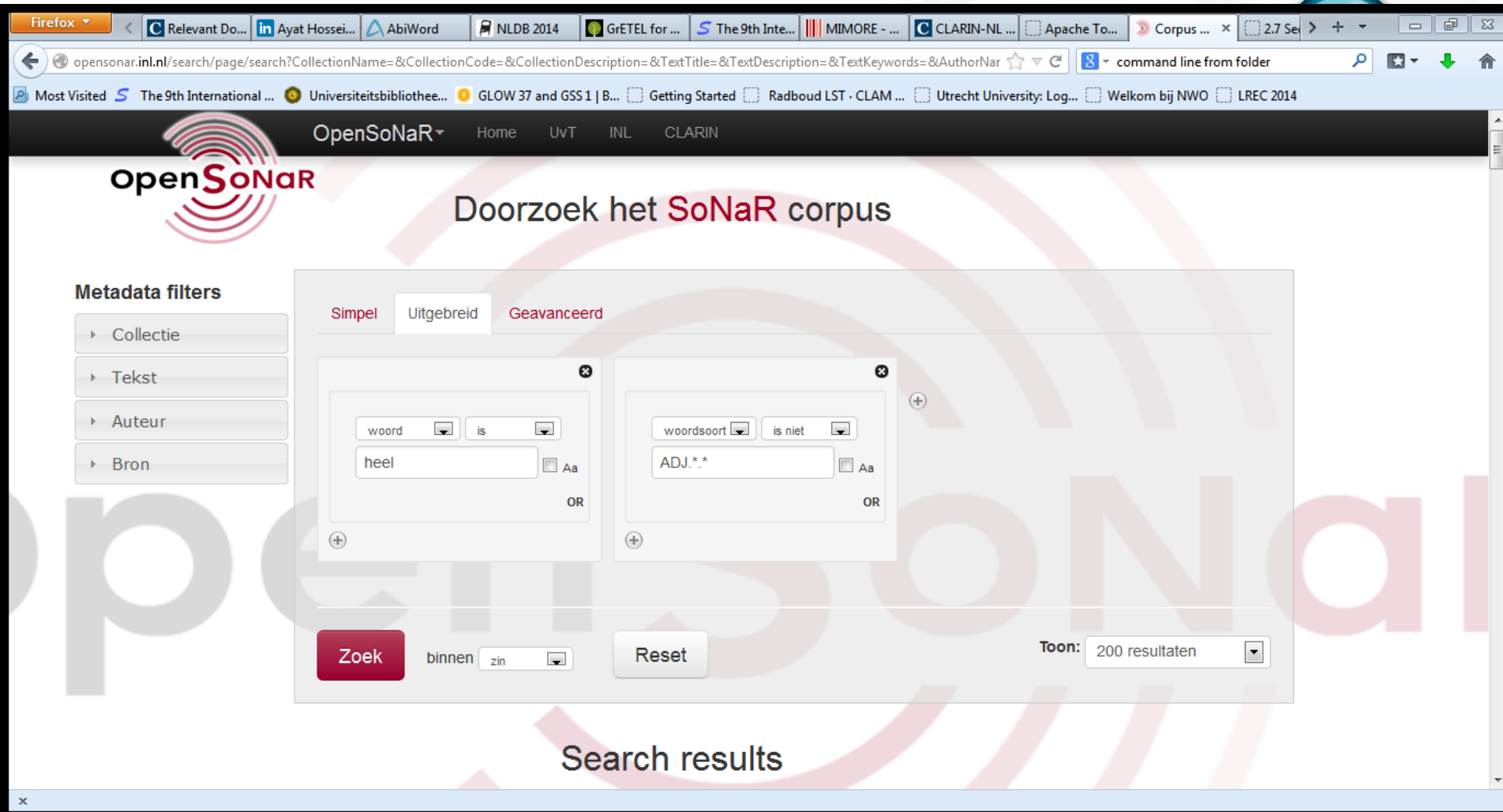
Most Visited | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS 1 | B... | Getting Started | Radboud LST - CLAM ... | Utrecht University: Log... | Welkom bij NWO | LREC 2014

OpenSoNAR | Home | UvT | INL | CLARIN

openSoNAR Total pages: 7.0

Prev | 1 | 4 | 5 | 6 | 7 | Next | Toggle titles

Left context ▾	Hit text	Right context ▾	Lemma	Part of speech
... dit alles . Je staat	<b>heel dichtbij</b>	... .. Gant . Wat ...	heel dichtbij	ADJ(vrij,basis,zonder) VZ(init)
... vrijheid . Goed om je	<b>heel te</b>	zien . Dit is Selena ...	heel te	ADJ(vrij,basis,zonder) VZ(init)
... . Je mag kiezen .	<b>Heel naar</b>	Japan , of in stukken ...	heel naar	ADJ(vrij,basis,zonder) VZ(init)
... enorm beïnvloed . Daar werd	<b>heel af</b>	en toe over gepraat . ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... Een lapsus . Serieus ,	<b>heel Ter</b>	Smissen laten opkalefateren door ons ...	heel ter	ADJ(vrij,basis,zonder) VZ(versm)
... van zijn vader stellen .	<b>Heel af</b>	en toe gaat Montgomery bij ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... het Belgische indoorkampioenschap is .	<b>Heel af</b>	en toe nog indoor-tornooien omdat ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... meer nodig om de robotjeep	<b>heel op</b>	Mars te krijgen . Het ...	heel op	ADJ(vrij,basis,zonder) VZ(init)
... op en spuugt je weer	<b>heel uit</b>	. Wil je op mij ...	heel uit	ADJ(vrij,basis,zonder) VZ(fin)
... . Er was niets meer	<b>heel aan</b>	die koe ! Het heeft ...	heel aan	ADJ(vrij,basis,zonder) VZ(init)
... eigenlijk heel vaak drama's .	<b>Heel af</b>	en toe een komedie omdat ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... mijn dochter niet zeggen .	<b>Heel af</b>	en toe eens een briefje ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... een echte steppevogel . Maar	<b>heel af</b>	en toe komt hij ook ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... Wat ik ga vertellen gebeurde	<b>heel in</b>	het begin van die heidense ...	heel in	ADJ(vrij,basis,zonder) VZ(init)
... het op seks aankomt .	<b>Heel af</b>	en toe is het wel ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... enorm beïnvloed . Daar werd	<b>heel af</b>	en toe over gepraat . ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)
... frequentie blijft open . Maar	<b>heel af</b>	en toe komen sommigen er ...	heel af	ADJ(vrij,basis,zonder) VZ(fin)



The screenshot shows the OpenSoNaR search interface in a Firefox browser. The browser's address bar shows the URL: `opensonar.in.nl/search/page/search?CollectionName=&CollectionCode=&CollectionDescription=&TextTitle=&TextDescription=&TextKeywords=&AuthorNar`. The page title is "Doorzoek het SoNaR corpus".

**OpenSoNaR** Home UvT INL CLARIN

## Doorzoek het SoNaR corpus

**Metadata filters**

- Collectie
- Tekst
- Auteur
- Bron

**Search filters:**

- Simpel** (selected) | Uitgebreid | Geavanceerd
- Filter 1: woord is heel (Aa) OR
- Filter 2: woordsoort is niet ADJ.\*\* (Aa) OR

**Search options:**

- Zoek** (button)
- binnen zin (dropdown)
- Reset (button)
- Toon: 200 resultaten (dropdown)

### Search results



Firefox | Relevant Do... | Ayat Hossei... | AbiWord | NLDB 2014 | GrETEL for ... | The 9th Inte... | MIMORE - ... | CLARIN-NL ... | Apache To... | Corpus ... x | 2.7 Se | + | - | x

opensonar.inl.nl/search/page/search?CollectionName=&CollectionCode=&CollectionDescription=&TextTitle=&TextDescription=&TextKeywords=&AuthorNar ☆ | & command line from folder | ☆ | ↓ | ↑

Most Visited | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS1 | B... | Getting Started | Radboud LST - CLAM ... | Utrecht University: Log... | Welkom bij NWO | LREC 2014

OpenSoNaR | Home | UvT | INL | CLARIN

**openSoNaR** respecten daarvan . **heel mijn** maag-darmsysteem was totaal overhoop . ... heel mijn ADJ(vrij,basis,zonder) VNW(bez,def,stan,vol,1,ev,prenom,zonder,agr)

... met christendom . Er wordt **heel weinig** over gepraat op dit ogenblik ... heel weinig ADJ(vrij,basis,zonder) VNW(onbep,grad,stan,vrij,zonder,basis)

... en aangenaam smaken . Nog **heel even** laten garen ... om ... heel even ADJ(vrij,basis,zonder) BW()

... originele windjammer ... in **heel de** wereld is . De visserij ... heel de ADJ(vrij,basis,zonder) LID(bep,stan,rest)

... de ezeltjes . Kinderen zijn **heel graag** met ezels bezig . Ezels ... heel graag ADJ(vrij,basis,zonder) BW()

... een boek met adressen over **heel de** wereld ... waar je ... heel de ADJ(vrij,basis,zonder) LID(bep,stan,rest)

... zeker . Vraag of hij **heel even** kan ... Rik ! Kom ... heel even ADJ(vrij,basis,zonder) BW()

... Theo Smit . Ik heb **heel wat** kritiek gekregen . Maar wat ... heel wat ADJ(vrij,basis,zonder) VNW(onbep,pron,stan,vol,3o,ev)

... openbare omroep . Ik heb **heel graag** dat beroep gedaan . Ik ... heel graag ADJ(vrij,basis,zonder) BW()

... komen . En hij heeft **heel zijn** leven moeten klimmen om hogerop ... heel zijn ADJ(vrij,basis,zonder) VNW(bez,def,stan,vol,3,ev,prenom,zonder,agr)

... gehad . Maar ik heb **heel ,** heel veel fantastische , mooie ... heel , ADJ(vrij,basis,zonder) LET()

... Maar ik heb heel , **heel veel** fantastische , mooie herinneringen . ... heel veel ADJ(vrij,basis,zonder) VNW(onbep,grad,stan,prenom,zonder,agr,basis)

... over . Want er zijn **heel wat** films die KNT zijn . ... heel wat ADJ(vrij,basis,zonder) VNW(onbep,pron,stan,vol,3o,ev)

Subtitles for deel 1 Hey . Welkom bij Mijn Gedacht . Gisteren is het 12e Europees Jeugdfilmfestival gestart . Alle films daar zijn kinderen toegelaten . Daar hebben we het vandaag over . Want er zijn **heel wat** films die KNT zijn . Kinderen Niet Toegelaten . Lauren ( 13 ) is een echte filmfreak . Ze gaat vaak naar de bioscoop . Af en toe mag ze niet binnen . Ik ben Lauren , 13 jaar . Soms ga ik 3 , 4 keer per maand naar

... chef . Je kunt niet **heel Scotsville** beoordelen op je medewerkers . ... heel Scotsville ADJ(vrij,basis,zonder) SPEC(deeleigen)

... op je lever ? - **Heel veel** . Maar ik wil graag ... heel veel ADJ(vrij,basis,zonder) VNW(onbep,grad,stan,vrij,zonder,basis)

# OPENSONAR

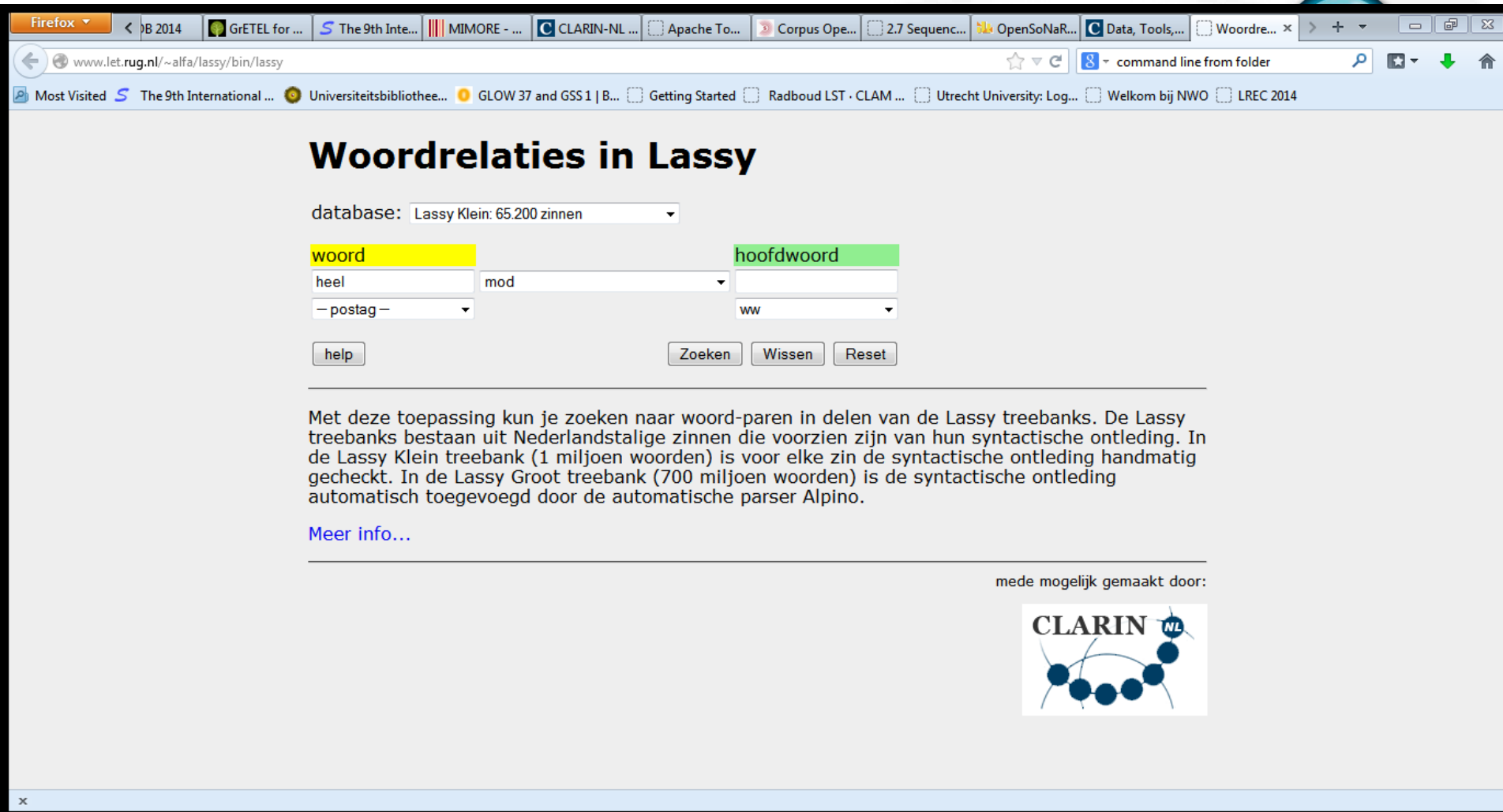


- [Return Page](#)

# LASSY SIMPLE INTERFACE

- Start Page





The screenshot shows a web browser window with the address bar displaying `www.let.rug.nl/~alfa/lassy/bin/lassy`. The page title is "Woordrelaties in Lassy". Below the title, there is a form for searching word pairs. The "database" dropdown is set to "Lassy Klein: 65.200 zinnen". The "woord" field contains "heel" and the "postag" dropdown is set to "- postag -". The "hoofdwoord" field is empty and the "ww" dropdown is set to "ww". There are three buttons: "help", "Zoeken", and "Wissen", and a "Reset" button. Below the form, there is a paragraph of text explaining the application's purpose and a link for "Meer info...". At the bottom right, it says "mede mogelijk gemaakt door:" followed by the CLARIN NL logo.

## Woordrelaties in Lassy

database: Lassy Klein: 65.200 zinnen

<b>woord</b>		<b>hoofdwoord</b>
heel	mod	
- postag -		ww


---

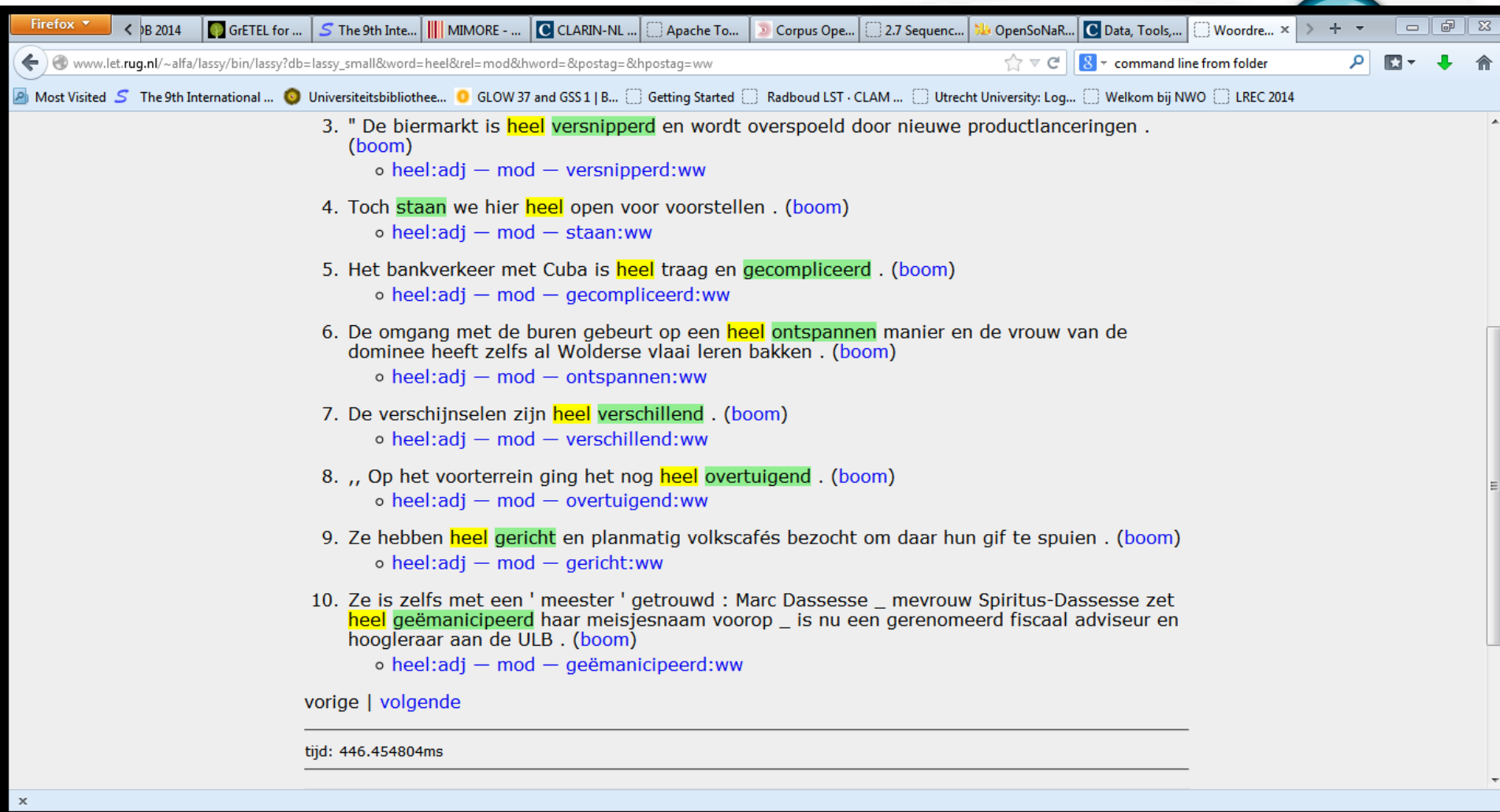
Met deze toepassing kun je zoeken naar woord-paren in delen van de Lassy treebanks. De Lassy treebanks bestaan uit Nederlandstalige zinnen die voorzien zijn van hun syntactische ontleding. In de Lassy Klein treebank (1 miljoen woorden) is voor elke zin de syntactische ontleding handmatig gecheckt. In de Lassy Groot treebank (700 miljoen woorden) is de syntactische ontleding automatisch toegevoegd door de automatische parser Alpino.

[Meer info...](#)

---

mede mogelijk gemaakt door:





Firefox < B 2014 GrETEL for ... The 9th Inte... MIMORE - ... CLARIN-NL ... Apache To... Corpus Ope... 2.7 Sequenc... OpenSoNaR... Data, Tools,... Woordre... x

www.let.rug.nl/~alfa/lassy/bin/lassy?db=lassy\_small&word=heel&rel=mod&hword=&postag=&hpostag=ww

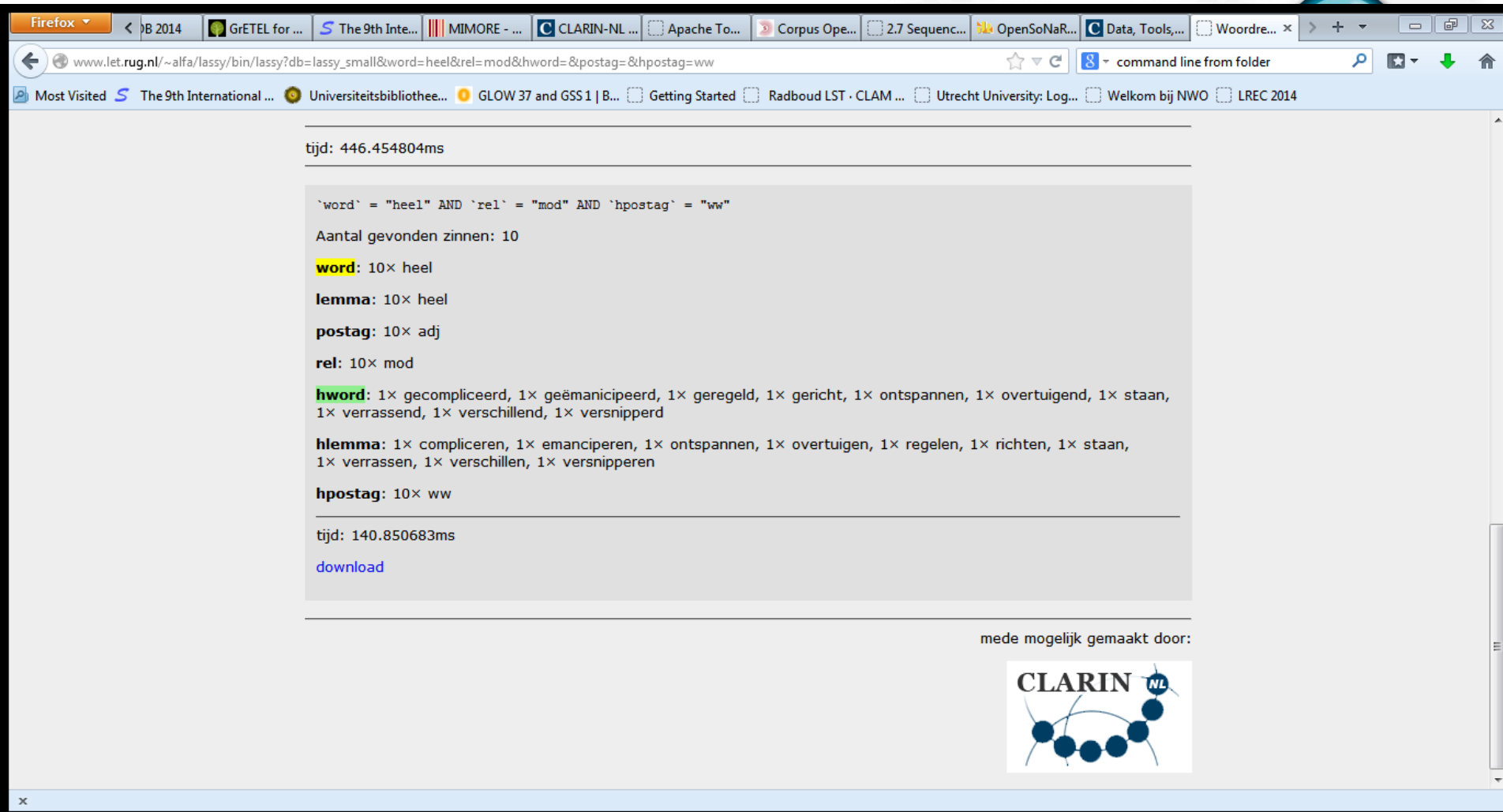
Most Visited The 9th International ... Universiteitsbibliothee... GLOW 37 and GSS1 | B... Getting Started Radboud LST · CLAM ... Utrecht University: Log... Welkom bij NWO LREC 2014

3. " De biermarkt is heel versnipperd en wordt overspoeld door nieuwe productlanceringen . (boom)
  - o heel:adj — mod — versnipperd:ww
4. Toch staan we hier heel open voor voorstellen . (boom)
  - o heel:adj — mod — staan:ww
5. Het bankverkeer met Cuba is heel traag en gecompliceerd . (boom)
  - o heel:adj — mod — gecompliceerd:ww
6. De omgang met de burens gebeurt op een heel ontspannen manier en de vrouw van de dominee heeft zelfs al Wolderse vlaai leren bakken . (boom)
  - o heel:adj — mod — ontspannen:ww
7. De verschijnselen zijn heel verschillend . (boom)
  - o heel:adj — mod — verschillend:ww
8. „ Op het voorterrein ging het nog heel overtuigend . (boom)
  - o heel:adj — mod — overtuigend:ww
9. Ze hebben heel gericht en planmatig volkscafés bezocht om daar hun gif te spuien . (boom)
  - o heel:adj — mod — gericht:ww
10. Ze is zelfs met een ' meester ' getrouwd : Marc Dassesse \_ mevrouw Spiritus-Dassesse zet heel geëmanicipeerd haar meisjesnaam voorop \_ is nu een gerenomeerd fiscaal adviseur en hoogleraar aan de ULB . (boom)
  - o heel:adj — mod — geëmanicipeerd:ww

vorige | volgende

---

tijd: 446.454804ms



tijd: 446.454804ms

---

```
`word` = "heel" AND `rel` = "mod" AND `hpostag` = "ww"
```

Aantal gevonden zinnen: 10

**word:** 10× heel

**lemma:** 10× heel

**postag:** 10× adj

**rel:** 10× mod

**hword:** 1× gecompliceerd, 1× geëmanicipeerd, 1× geregeld, 1× gericht, 1× ontspannen, 1× overtuigend, 1× staan, 1× verrassend, 1× verschillend, 1× versnipperd

**hlemma:** 1× compliceren, 1× emanciperen, 1× ontspannen, 1× overtuigen, 1× regelen, 1× richten, 1× staan, 1× verrassen, 1× verschillen, 1× versnipperen

**hpostag:** 10× ww


---

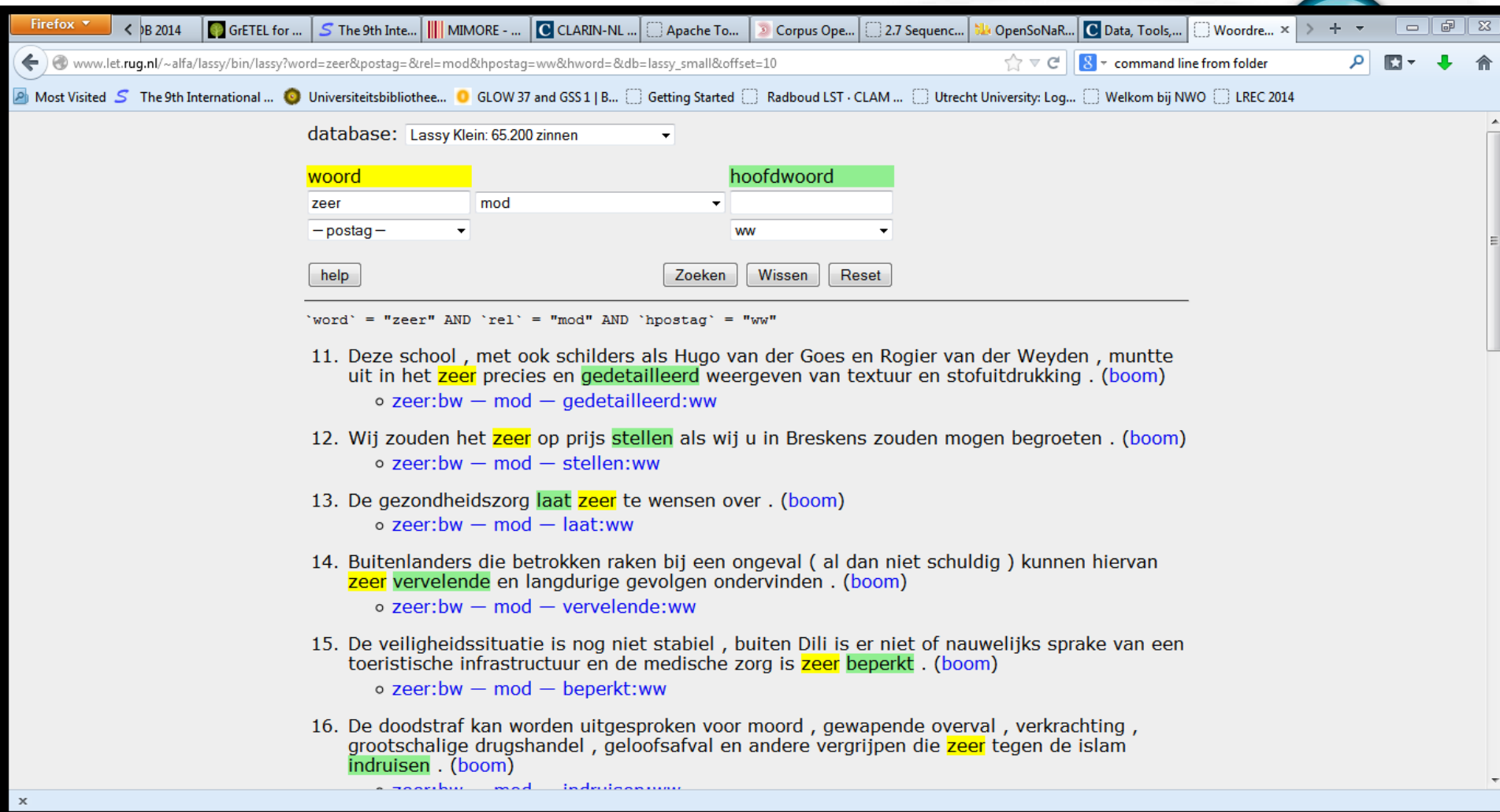
tijd: 140.850683ms

[download](#)

---

mede mogelijk gemaakt door:





database: Lassy Klein: 65.200 zinnen

woord: zeer    mod: mod    hoofdwoord:    hpostag: ww

help    Zoeken    Wissen    Reset

`word` = "zeer" AND `rel` = "mod" AND `hpostag` = "ww"

11. Deze school , met ook schilders als Hugo van der Goes en Rogier van der Weyden , muntte uit in het zeer precies en gedetailleerd weergeven van textuur en stofuitdrukking . (boom)
  - o zeer:bw — mod — gedetailleerd:ww
12. Wij zouden het zeer op prijs stellen als wij u in Breskens zouden mogen begroeten . (boom)
  - o zeer:bw — mod — stellen:ww
13. De gezondheidszorg laat zeer te wensen over . (boom)
  - o zeer:bw — mod — laat:ww
14. Buitenlanders die betrokken raken bij een ongeval ( al dan niet schuldig ) kunnen hiervan zeer vervelende en langdurige gevolgen ondervinden . (boom)
  - o zeer:bw — mod — vervelende:ww
15. De veiligheidssituatie is nog niet stabiel , buiten Dili is er niet of nauwelijks sprake van een toeristische infrastructuur en de medische zorg is zeer beperkt . (boom)
  - o zeer:bw — mod — beperkt:ww
16. De doodstraf kan worden uitgesproken voor moord , gewapende overval , verkrachting , grootschalige drugshandel , geloofsafval en andere vergrijpen die zeer tegen de islam indruisen . (boom)
  - o zeer:bw — mod — indruisen:ww

Firefox < B 2014 GrEtel for ... The 9th Inte... MIMORE - ... CLARIN-NL ... Apache To... Corpus Ope... 2.7 Sequenc... OpenSoNaR... Data, Tools,... Woordre... x

www.let.rug.nl/~alfa/lassy/bin/lassy?word=zeer&postag=&rel=mod&hpostag=ww&hword=&db=lassy\_small&offset=10

Most Visited The 9th International ... Universiteitsbibliothee... GLOW 37 and GSS1 | B... Getting Started Radboud LST · CLAM ... Utrecht University: Log... Welkom bij NWO LREC 2014

[vorige](#) | [volgende](#)

---

tijd: 182.203374ms

---

```
`word` = "zeer" AND `rel` = "mod" AND `hpostag` = "ww"
```

Aantal gevonden zinnen: 51

**word:** 49× zeer, 1× zéér, 1× Zeer

**lemma:** 51× zeer

**postag:** 51× bw

**rel:** 51× mod

**hword:** 4× geïnteresseerd, 3× beperkt, 2× gedetailleerd, 2× ingenomen, 2× is, 2× teleurgesteld, 2× vervelende, 1× behoudende, 1× bemoedigende, 1× beperkte, 1× betreuren, 1× gecompliceerd, 1× gedetailleerde, 1× geïrriteerd, 1× gematigd, 1× gemengde, 1× gemotiveerd, 1× geperfectioneerde, 1× geregeld, 1× gerespecteerde, 1× geslaagd, 1× gewaardeerd, 1× gewild, 1× gewilde, 1× indruisen, 1× ingrijpend, 1× invaliderende, 1× laat, 1× lijden, 1× maakte, 1× ontspannen, 1× opvallend, 1× overtreft, 1× stellen, 1× uitgebreide, 1× verdedigende, 1× verheugen, 1× verontrust, 1× verontwaardigd, 1× verrassend, 1× verschillend

**hlemma:** 4× beperken, 4× interesseren, 3× detailleren, 2× innemen, 2× teleurstellen, 2× vervelen, 2× willen, 2× zijn, 1× behouden, 1× bemoedigen, 1× betreuren, 1× compliceren, 1× indruisen, 1× ingrijpen, 1× invalideren, 1× irriteren, 1× laten, 1× lijden, 1× maken, 1× matigen, 1× mengen, 1× motiveren, 1× ontspannen, 1× opvallen, 1× overtreffen, 1× perfectioneren, 1× regelen, 1× respecteren, 1× slagen, 1× stellen, 1× uitbreiden, 1× verdedigen, 1× verheugen, 1× verontrusten, 1× verontwaardigen, 1× verrassen, 1× verschillen, 1× waarden

**hpostag:** 51× ww

---

tijd: 76.789204ms

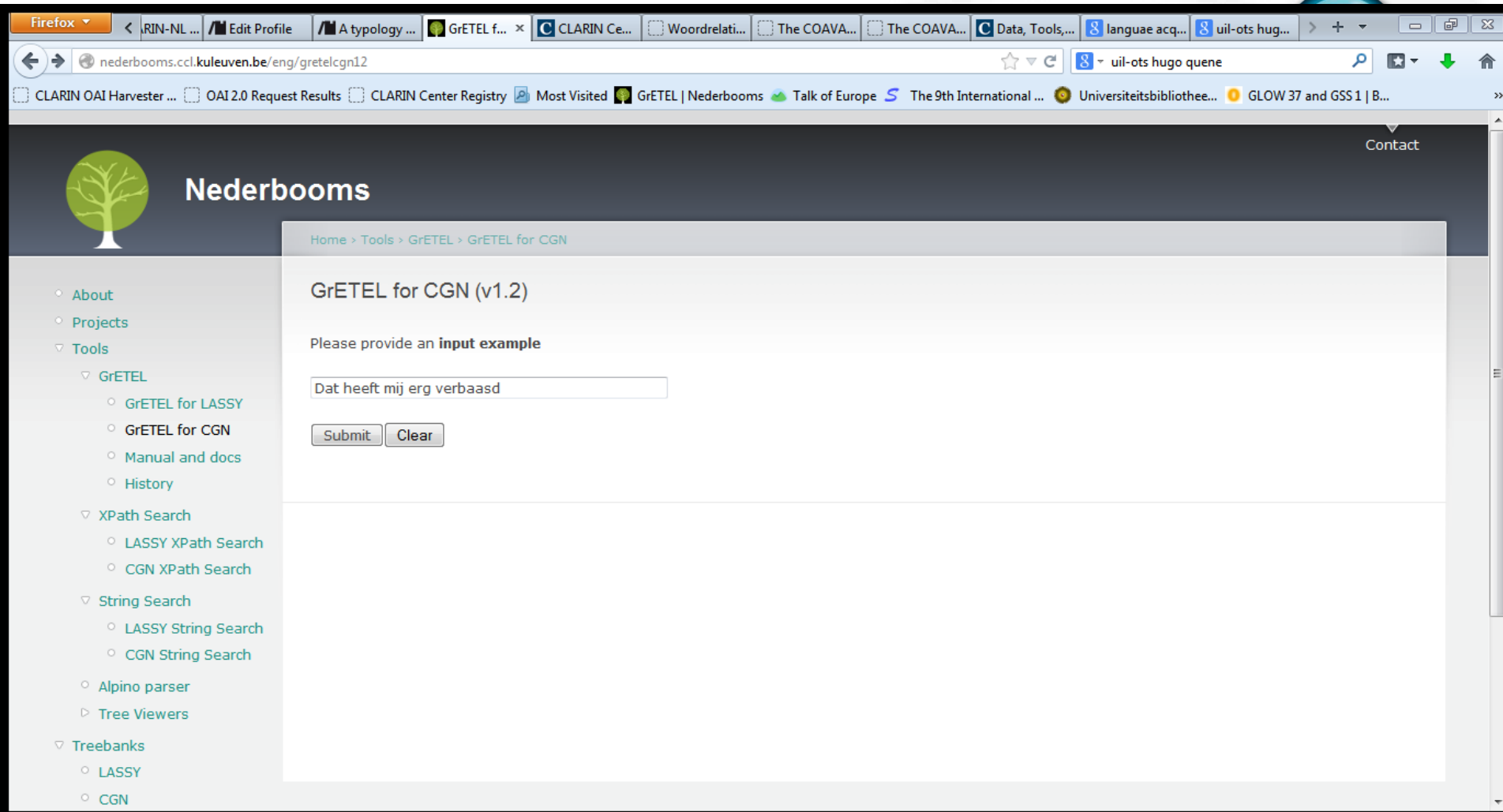
[download](#)



# LASSY SIMPLE INTERFACE



- Return Page



The screenshot shows a web browser window displaying the GreTEL for CGN application. The browser's address bar shows the URL `nederbooms.ccl.kuleuven.be/eng/gretelcg12`. The page header includes the "Nederbooms" logo (a green tree) and a navigation menu with items like "About", "Projects", "Tools", "XPath Search", "String Search", "Alpino parser", "Tree Viewers", and "Treebanks". The main content area is titled "GreTEL for CGN (v1.2)" and contains a form for providing an input example. The form has a text input field with the value "Dat heeft mij erg verbaasd" and two buttons labeled "Submit" and "Clear". A breadcrumb trail at the top of the content area reads "Home > Tools > GrETEL > GrETEL for CGN".

Home > Tools > GrETEL > GrETEL for CGN

## GrETEL for CGN (v1.2)

Please provide an **input example**

Firefox | RIN-NL ... | Edit Profile | A typology ... | GrETEL f... x | CLARIN Ce... | Woordrelati... | The COAVA... | The COAVA... | Data, Tools,... | languae acq... | uil-ots hug... | uil-ots hugo quene

nederbooms.ccl.kuleuven.be/eng/node/54

CLARIN OAI Harvester ... | OAI 2.0 Request Results | CLARIN Center Registry | Most Visited | GrETEL | Nederbooms | Talk of Europe | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS 1 | B...

sentence		Dat	heeft	mij	erg	verbaasd
relevant nodes	pos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
	extended pos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	lemma	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
	token	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
optional nodes		<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

**OPTIONS**

Respect word order  
 Ignore properties of the dominating node  
 Split extended pos tags

**GUIDELINES**

- pos**: Short part-of-speech tag (e.g. *n*, *ww*, *vnw*). [\[list of all pos tags\]](#)
- extended pos**: Long part-of-speech tag. For example: *N(soort,mv,basis)*, *WW(pv,tgw,ev)*, *VNW(pers,pron,nomin,vol,2v,ev)*. [\[list of all pos tags\]](#)
- lemma**: Word form that generalizes over inflected forms. For example: *zin* is the lemma of *zin*, *zinnen*, and *zinnetje*; *gaan* is the lemma of *ga*, *gaat*, *gaan*, *ging*, *gingen*, and *gegaan*. Lemma is case insensitive (except for proper names).
- token**: The exact word form. This is a case sensitive feature.

Treebank		Contents	# Sentences	# Words	# Sentences	# Words	# Sentences	# Words
<input checked="" type="checkbox"/> NL	<input checked="" type="checkbox"/> VL		NL	VL			TOTAL	
<input checked="" type="checkbox"/> NA	<input checked="" type="checkbox"/> VA	Spontaneous conversations ('face-to-face')	50,239	302,828	22,881	147,418	73,120	450,246
<input checked="" type="checkbox"/> NB	<input checked="" type="checkbox"/> VB	Interviews with teachers of Dutch	2,484	25,724	4,289	34,158	6,773	59,882
<input checked="" type="checkbox"/> NC	<input checked="" type="checkbox"/> VC	Telephone conversations (recorded via a switchboard)	11,649	70,084	3,142	19,984	14,791	90,068
<input type="checkbox"/> ND	<input checked="" type="checkbox"/> VD	Telephone conversations (recorded on MD)	0	0	929	6,309	929	6,309
<input checked="" type="checkbox"/> NE	<input type="checkbox"/> VE	Simulated business negotiations	3,123	25,524	0	0	3,123	25,524
<input checked="" type="checkbox"/> NF	<input checked="" type="checkbox"/> VF	Interviews/discussions/debates (broadcast)	6,290	75,167	2,617	25,122	8,907	100,289
<input checked="" type="checkbox"/> NG	<input checked="" type="checkbox"/> VG	(Political) discussions/debates/meetings (non-broadcast)	1,166	25,125	543	9,009	1,709	34,134
<input checked="" type="checkbox"/> NH	<input checked="" type="checkbox"/> VH	Lessons recorded in the classroom	3,064	26,004	1,395	10,116	4,459	36,120
<input checked="" type="checkbox"/> NI	<input checked="" type="checkbox"/> VI	Live (sports) commentaries (broadcast)	2,251	25,002	1,026	10,147	3,277	35,149
<input checked="" type="checkbox"/> NJ	<input checked="" type="checkbox"/> VJ	Newsreports (broadcast)	2,259	25,084	536	7,686	2,795	32,770
<input checked="" type="checkbox"/> NK	<input checked="" type="checkbox"/> VK	News (broadcast)	1,923	25,353	558	7,306	2,481	32,659
<input checked="" type="checkbox"/> NL	<input checked="" type="checkbox"/> VL	Commentaries/columns/reviews (broadcast)	1,857	25,082	601	7,431	2,458	32,513
<input checked="" type="checkbox"/> NM	<input checked="" type="checkbox"/> VM	Ceremonious speeches/sermons	444	5,190	107	1,894	551	7,084
<input checked="" type="checkbox"/> NN	<input checked="" type="checkbox"/> VN	Lectures/seminars	593	14,921	701	8,159	1,294	23,080
<input type="checkbox"/> NO	<input checked="" type="checkbox"/> VO	Read speech	0	0	3,256	44,144	3,256	44,144
<b>CGN core</b>		Complete treebank	87,342	671,088	42,581	338,883	129,923	1,009,971

**OPTION**

Include context (one sentence before and after the matching sentence)

Firefox | RIN-NL ... | Edit Profile | A typology ... | GrETEL f... x | CLARIN Ce... | Woordrelati... | The COAVA... | The COAVA... | Data, Tools,... | langue acq... | uil-ots hug... | + | - | x

nederbooms.ccl.kuleuven.be/eng/node/56 | uil-ots hugo quene

CLARIN OAI Harvester ... | OAI 2.0 Request Results | CLARIN Center Registry | Most Visited | GrETEL | Nederbooms | Talk of Europe | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS1 | B...

Download matches[ext] | Printer-friendly version

Search within results:

SENTENCE ID	MATCHING SENTENCE	HITS	DISPLAY OPTIONS
<a href="#">fnf007277__28</a>	nou 't varieert erg per traject .	1	[full screen] [XML]
<a href="#">fnj007333__181</a>	ik moet d'r vaak erg om lachen maar dan ...	1	[full screen] [XML]
<a href="#">fvb400169__219</a>	want dus andere richtingen bijvoorbeeld handel talen en zo <b>daar ligt de klemtoon dan natuurlijk wel erg op de talen</b> .	1	[full screen] [XML]
<a href="#">fna000432__56</a>	oei <b>dat ging wel erg</b> ...	1	[full screen] [XML]
<a href="#">fni007174__85</a>	<b>deed me erg denken aan de provocatie van Vercouteren tegen Wim Kieft ooit</b> .	1	[full screen] [XML]
<a href="#">fvf600990__52</a>	er zijn Vlaams-nationale teksten over de joden beschikbaar dat weet u ook <b>die erg lijken op die van het Vlaams Blok over de islamieten van vandaag</b> .	1	[full screen] [XML]
<a href="#">fna000250__214</a>	nee dan moet je al erg aan de <b>aan de rand</b> uh gaan zitten .	1	[full screen] [XML]
<a href="#">fnb000089__7</a>	da 's natuurlijk een <b>erg afgezaagd</b> antwoord maar dat was 't wel .	1	[full screen] [XML]

SENTENCE ID	MATCHING SENTENCE	HITS	DISPLAY OPTIONS
-------------	-------------------	------	-----------------

Showing 1 to 36 of 36 entries

# GRETEL CGN



- [Return Page](#)

Firefox | TEL for ... | Data, Tools, ... | Zoekresu... x | Arthurian Fi... | Woordrelati... | The COAVA... | The COAVA... | CLARIN Virt... | New Tab | How do I p... | Using Large... | cornetto.inl.nl/cornetto/cornetto\_simple\_search.xml | textcavator

CLARIN OAI Harvester ... | OAI 2.0 Request Results | CLARIN Center Registry | Most Visited | GrEtel | Nederbooms | Talk of Europe | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS1 | B...

## Cornetto Simple Search

Word form:  POS:  Items per page:

Use regular expressions  ?

### Matching lexical entries

selection (max. 500) 1 to 6 from 6

ID	Written Form	Semantics	POS	Examples	Pragmatics	Syntax
zeer-a-1	zeer	pijnlijk	adjective	een zere knie/arm (1 more)		
zeer-adv-3	zeer	(no semantics)	adverb			
zeer-mwe-83673	op zijn zeer trappen	spreken over zaken die iemand kwetsen				
zeer-mwe-83675	iemand in zijn zeer tasten	spreken over zaken die iemand kwetsen				
zeer-n-1	zeer	pijn	noun	zeer doen (2 more)		
zeer-n-2	zeer	zere plek	noun	kwaad zeer		

### Details of selected lexical entry (zeer-adv-3)

General	Relations	Hierarchy	Word forms
<b>lemma:</b> zeer <b>part of speech:</b> adverb <b>form type:</b> <b>mode:</b> <b>morphology:</b> <b>morpho-syntax:</b>	<b>type lexical entry</b> (No sense relations) <b>var. lemma type</b> (No form relations)	<a href="#">nld-21-n r-535423-r</a> <a href="#">allemachtig-adv-2</a> <a href="#">beestachtig-adv-2</a> <a href="#">bijzonder-a-4</a> <a href="#">bliksems-adv-2</a> <a href="#">bloedig-adv-2</a> <a href="#">bovenmate-adv-1</a> <a href="#">buitengewoon-adv-2</a> <a href="#">buitenmate-adv-1</a> <a href="#">buitensporig-adv-2</a> <a href="#">crimineel-a-4</a> <a href="#">deerlijk-adv-2</a> <a href="#">deksels-adv-2</a> <a href="#">donders-adv-2</a> <a href="#">drommels-adv-2</a> <a href="#">eindeloos-a-3</a> <a href="#">enorm-adv-2</a> <a href="#">erbarmelijk-adv-2</a> <a href="#">fantastisch-adv-6</a> <a href="#">formidabel-adv-2</a> <a href="#">geweldig-adv-4</a> <a href="#">goddeloos-adv-2</a> <a href="#">godsjammerlijk-adv-2</a> <a href="#">grenzeloos-adv-2</a> <a href="#">grotelijks-adv-1</a> <a href="#">heel-adv-5</a> <a href="#">ijselijk-adv-2</a> <a href="#">ijzig-a-4</a> <a href="#">intens-adv-2</a> <a href="#">krankzinnig-adv-3</a> <a href="#">machtig-adv-4</a> <a href="#">mirakels-adv-1</a> <a href="#">monsterachtig-adv-2</a> <a href="#">moorddadig-adv-4</a> <a href="#">oneindig-adv-2</a> <a href="#">onnoemelijk-adv-2</a> <a href="#">ontiegelijk-adv-2</a> <a href="#">ontstellend-adv-2</a> <a href="#">ontzaggelijk-adv-2</a> <a href="#">ontzettend-adv-3</a> <a href="#">onuitsprekelijk-adv-2</a> <a href="#">onvoorstelbaar-adv-2</a> <a href="#">onwezenlijk-adv-2</a> <a href="#">onwijs-adv-4</a> <a href="#">overweldigend-adv-2</a> <a href="#">peilloos-adv-2</a> <a href="#">reusachtig-adv-3</a> <a href="#">reuze-adv-2</a> <a href="#">schrikkelijk-adv-2</a> <a href="#">sterk-adv-7</a> <a href="#">uiterst-adv-4</a> <a href="#">verdomd-adv-2</a> <a href="#">verdraaid-a-4</a> <a href="#">verduiveld-adv-2</a> <a href="#">verduveld-adv-2</a> <a href="#">verrekt-adv-3</a> <a href="#">verrot-adv-3</a> <a href="#">verschrikkelijk-adv-3</a> <a href="#">vervloekt-adv-2</a> <a href="#">vreselijk-adv-5</a> <a href="#">waaninnig-adv-2</a> <a href="#">zeer-adv-3</a> <a href="#">zeldzaam-adv-2</a> <a href="#">zwaar-adv-10</a>	<b>form</b> zeer
<b>Sense</b> <b>definition:</b> <b>polarity:</b> negative	<b>Semantics</b> <b>semantic type:</b> <b>semantic shift:</b>	<b>Examples</b> (No examples)	<b>Syntax</b> (No syntactics)

Firefox | TEL for ... | Data, Tools, ... | Zoekresu... x | Arthurian Fi... | Woordrelati... | The COAVA... | The COAVA... | CLARIN Virt... | New Tab | How do I p... | Using Large... | cornetto.inl.nl/cornetto/cornetto\_simple\_search.xql | textcavator

CLARIN OAI Harvester ... | OAI2.0 Request Results | CLARIN Center Registry | Most Visited | GrEtel | Nederbooms | Talk of Europe | The 9th International ... | Universiteitsbibliothee... | GLOW 37 and GSS1 | B...

## Cornetto Simple Search

Word form:  POS:  Items per page:

Use regular expressions  ?

**Matching lexical entries** [help](#) [home](#)

selection (max. 500) 1 to 12 from 12

ID	Written Form	Semantics	POS	Examples	Pragmatics	Syntax
erg-a-1	erg	naar	adjective	iets erg vinden (5 more)		
erg-a-2	erg	ontzettend	adjective	erge honger/dorst/slaap/haast hebben (3 more)		
erg-mwe-85465	geen erg hebben in iets	het niet merken				
erg-mwe-85466	zonder erg	onopzettelijk				
erg-mwe-r_a-10434-11	het ergste hebben we gehad	is achter de rug				
erg-mwe-r_a-10434-12	op het ergste voorbereid zijn	uitgaan van, rekening houden met het meest ongunstige geval				
erg-mwe-r_a-10434-13	zo erg is het niet	het valt wel mee				
erg-mwe-r_a-10434-2	iets (nog) erger maken dan het (al) is	het overdrijven		je moet de dingen niet erger maken dan ze zijn		
erg-mwe-r_a-10434-6	in het ergste geval	(no semantics)				
erg-mwe-r_a-10434-7	het is niet erg	het geeft niet				
erg-mwe-r_a-10434-9	het middel is erger dan de kwaal	(no semantics)				
erg-n-1	erg	kwade bedoeling	noun			

### Details of selected lexical entry (erg-a-2)

General	Relations	Hierarchy	Word forms
lemma: erg	type lexical entry	nld-21- n a-512478-a	form comp
part of speech: adjective	(No sense relations)	erg-a-2 ernstig-a-2 fel-a-1 hard-a-4 heftig-a-1 hevig-a-1 krachtig-a-3 sterk-a-4 stevig-a-2 straf-a-2 vet-a-5 vurig-a-1 zwaar-a-3	erg erger comparative





- [Return Page](#)

Firefox | yago.meertens.knaw.nl/CoavaMainApplication/ | nederbooms gretel

CLARIN OAI Harvester ... OAI 2.0 Request Results ... CLARIN Center Registry ... Most Visited ... GrETEL | Nederbooms ... Talk of Europe ... The 9th International ... Universiteitsbibliothee... GLOW 37 and GSS1 | B...

**Actions**

Please choose a view  
 Focus on a group of children

Add words :  
 suggest terms/freq  
 quantity (max): 15  
 Hint  
 clear

Select one or more corpora :  
 antwerp   
 bol   
 gillis   
 groningen   
 schaarlaekens   
 vankampen   
 wijnen   
 deselect all

Choose a age step size :  
 3 months

Adjust the offsets of the age axis :  
 500 - 2100

**Correlation between word count and child age in days**

Age in days	zeer	erg	heel
591	2	0	0
773	5	0	2
955	5	0	25
1137	15	5	70
1319	5	5	85
1501	5	5	140
1683	5	5	50
1865	5	5	45
2047	5	5	25

- [Return Page](#)



# GRETEL CGN



- [Return Page](#)

# OTHER EXAMPLES



- *PP/A*
  - *In zijn sas, in verwachting, tegen, voor, onder de indruk, uit de tijd*
  - *Tevreden met v. in zijn sas met*
  - *Zwanger v. in verwachting*
  - *Verward v. in de war*
  - *Modieus v. in de mode / in zwang*
- English: *very v. very much*
- *V:*
  - *Worden (AP, NP, \*PP) v. raken (AP, \*NP, PP)*

# CHILD-DIRECTED SPEECH



- *Heel, zeer, erg* in children-addressed speech (Van Kampen only):

	Mod A	Mod N	Mod V	Mod P	Pred	Other	Unclear
heel	421	10	2	0	7	1	4
erg	2	0	2	0	37	0	0
zeer	33	2	0	0	54	0	2