CLARIN-*NL* Annual Report 2015

1   Introduction

This document provides a description of the activities carried out in CLARIN-NL for 2015. Explanations and expansions of acronyms in this document can be found on the CLARIN-NL website[1] and in a separate document.

2   Executive Summary

Through an extension awarded by NWO the CLARIN-NL project ran during the whole of 2015 instead of until April 1st, 2015 only. We summarize the major highlights here.

A successful and well-visited final event was organized on March 13, 2015 at the premises of NISV in combination with the kick-off of the CLARIAH-CORE project. Many demonstrations of applications developed and data curated in the CLARIN-NL project were given there.

After the extension was awarded, five small projects were started to extend work initiated earlier in CLARIN-NL, on new functionality made necessary by the changing role of MPI/TLA, and to connect CLARIN with Taalportaal.

The Netherlands was well represented at the annual CLARIN Conference (Wrocław, Poland), and Maarten van Gompel was awarded the CLARIN Young Scientist Award there.

It is clear that the CLARIN infrastructure is increasingly used in research. Research that crucially used the CLARIN infrastructure is now appearing in the *Lingua* Special Issue on the CLARIN Infrastructure with Jan Odijk as guest editor. We are approached by many researchers when they are preparing new project proposals. CLARIN is also involved in a project proposal for the NWO *Gravitation* programme that is being prepared by a large consortium of linguists united in LOT. The first PhD in which crucial use is made of CLARIN has been successfully defended by Liesbeth Augustinus in Leuven. The CLARIN work for this was made possible in part through the close collaboration of Flanders with CLARIN-NL.

There have been many educational activities, including a LOT Winterschool course and multiple lectures on specific aspects of CLARIN at least in Utrecht and in Amsterdam (UvA). Some of these have become part of the regular curriculum.

There have been many dissemination actions, inter alia at the TIN conference (Utrecht),the CLIN conferences (Antwerp and Amsterdam), at the DRONGO festival (Utrecht), and internationally at the Austrian Academy of Sciences (Vienna, Austria), at the Digital Humanities Spring School (Leuven, Belgium), and at the PARSEME workshop (Iaşi, Romania) . The CLARIN-TCC project held two

---

[1] http://www.clarin.nl/node/8

successful Creative Camps (The Hague and Amsterdam) and is now preparing a European research proposal around parliamentary data in Europe using the CLARIN infrastructure. Three new movies on specific data sets and applications were produced. A book on CLARIN-NL (editors: Jan Odijk and Arjan van Hessen) with 26 submitted contributions is in preparation and its Open Access publication is expected by the end of 2016.

3   Planned Activities

The original end date of the CLARIN-NL project was April 1st, 2015. An extension until the end of 2015 was requested and awarded. This made it possible to start up a few small projects in 2015.

1.1   EU-Line (WBS 1.9)

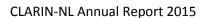No budget for this item was foreseen or spent.

Alice Dijkstra is the representative for the Netherlands in the CLARIN ERIC General Assembly, with Jan Odijk as expert.

CLARIN-NL is very active with regard to the European enterprise to create a CLARIN infrastructure under the guidance of CLARIN ERIC, and in other CLARIN or infrastructure-related projects and initiatives. This can be illustrated with many examples:

- CLARIN ERIC is hosted by the Netherlands, more specifically Utrecht University
- The CLARIN ERIC director was Steven Krauwer, and currently is Franciska de Jong.
- NL is represented in the National Coordinator's Forum, attends the monthly (virtual) meetings, and is active in it. (Jan Odijk/Daan Broeder)
- Active participation in CLARIN ERIC committees, e.g.
    - o   Marc Kemps-Snijders, Member Standing Committee for CLARIN Technical Centres
    - o   Daan Broeder member Centre Assessment Committee
    - o   Daan Broeder, member CLARIN ERIC standards committee
    - o   Menzo Windhouwer, member CLARIN ERIC standards committee
- A large delegation from the Netherlands attended the CLARIN Annual Conference in Wrocław, Poland, and many of them had contributions in the form of presentations and/or showing demonstrators (about 10 presentations, demos or posters with NL contributors).
- Many CLARIN-NL people and/or organizations are involved in important European projects and initiatives that are directly relevant to infrastructures in general and the CLARIN infrastructure in particular. The Netherlands is involved in 3 out of the 5 currently running European projects that CLARIN ERIC participates in:
    - o   CLARIN PLUS: Meertens Institute
    - o   PARTHENOS: Meertens Institute via secondment
    - o   EUDAT2020: DANS, and Meertens Institute via secondment

(the two other projects are LT-Observatory and Europeana-DSI)

1.2 NL-Line

1.2.1 Technical (WBS 1.1)

Since the extension of the project to Dec 31, 2015 was known rather late, most activities in the period from Jan 1, 2015 through March 31, 2015 were focused on rounding off the CLARIN-NL project and smoothing the transition to the successor project CLARIAH-CORE. On the technical side, these include especially work on integration via the CLARIN Portal, some work on federated login and work on semantic interoperability.

1.2.2 Integration

An important activity started up in 2013 was an *integration activity (*the *CLAPOP project)* in which the various components created separately in a variety of projects (infrastructure functionality, search functionality, curated data, curated tools, demonstrators, web applications, etc. etc.) come together and are made accessible via a single portal to the CLARIN infrastructure. The CLARIN-NL portal has been created, and it contains

- Faceted search for services (tools, web applications, web services), which include some properties which were proposed in the *Metadata for Tools project* as facets
- Faceted Search for data
- Overview of the Dutch CLARIN Centres
- Links to metadata search applications and to federated content search
- Overview of educational packages developed in CLARIN-NL and other educational material
- An initial version of registry and faceted browsing for standards and best practices supported by CLARIN. This work was meant as a show-model to discuss a standardisation registry with CLARIN EU and other CLARIN projects.

The portal uses the CLARIN federated identity management setup to avoid multiple authentications when already logged-in with a CLARIN service.

The idea is that the portal will continue to exist also long after the CLARIN-NL project has finished. Its maintenance will be dealt with in CLARIAH after the end of the CLARIN-NL project.

Though it currently is a CLARIN-NL portal, it should turn into a portal containing all information that is relevant to CLARIN in the Netherlands, and perhaps even, in the context of CLARIAH, to a Digital Humanities portal. The services registry and faceted search were quite successful, and have been adopted by CLARIN Poland, and discussions are on-going with CLARIAH Austria and LINDAT (Czech Republic) on how to raise it to the EU-level.

1.2.3 Federated Login

Federated login was extended to a small number of applications that still lacked this.

1.2.4 Data, Tools and Services (WBS 1.2/1.3/1.4)

A number of new projects, small in scale, were started up in 2015. The projects were suggested by the executive board members or based on suggestions by other participants in CLARIN-NL. The EB made a proposal for starting up 5 new projects. This proposal was approved by the board (via e-mail)

in May 2015. The selected projects add important data and functionality to the CLARIN infrastructure:

- The Meertens Deposition Facility project lays the foundation for a solid deposition facility at Meertens, which became rather urgent given the reduced activities of the Max Planck Institute, and the uncertainty on easy accessibility of relevant data via MPG in the long run.
- The CARE project continues data curation activities on the regional dictionaries, making them available as databases and with associated CLARIN-compatible CMDI-metadata. Parts of some of these dictionaries were curated in the earlier COAVA project, but CARE will make more parts of more of these dictionaries available.
- The CorpusStudioWeb project turns a stand-alone application for search and analysis of linguistically annotated corpora into a web application, thus making it available and usable by a far larger group of researchers. The CorpusStudio application had proven its usefulness already by its wide use by researchers, especially in the context of research into historical linguistics.
- The OpenCGN project turns the OpenSONAR web application into a web application that allows combined search in the SONAR Corpus of written Dutch and the CGN Corpus of Spoken Dutch.
- The TPC project creates links from Taalportaal to the front-end of search and analysis applications operating on grammatical corpora and databases, e.g. to illustrate syntactic constructions (via search in syntactically annotated corpora), morphological, and phonological phenomena (via links to the CELEX database). It thus enriches the Taalportaal, stimulates the use of corpora and lexical databases by linguists in their research and makes linguists gradually acquainted with such corpus search applications.

It was also planned to transition the Groningen PaQu application to INL so that it would be running on the servers of a recognized CLARIN Centre. However, since a lot of new development work on PaQu was started up in 2015 in the context of the CLARIAH-CORE project, it was decided to postpone this to later as a part of the CLARIAH-CORE project.

1.2.5   2015 Projects

We briefly state the major properties of the projects started in 2015.

CARE	CuRation and integration of REgional dictionaries
Project Coordinator	Prof.dr. Nicoline van der Sijs (Nijmegen University)
Goal	Curation of various regional dictionaries of the Dutch language. Turning them into lexical databases. The project also lays a basis for connecting these databases to other regional databases such as eWND, DynaSand, and MIMORE.
Budget	€ 52.8k
Centre	Meertens Institute
Scientific impact	The curation of these regional dictionaries is crucial for the study of Dutch dialects and language variation in general.
Status	The project has effectively finished and has submitted its final reports, which have been approved. The actual data and metadata are expected to become available through the CLARIN VLO in June 2016.

CorpusSearchWeb	Corpus Studio – web version
Project Coordinator	Dr. Erwin Komen (Meertens Institute)
Goal	The project aims to make a web version of the "Corpus Studio" programme. The web version will be a reduced version of the stand-alone version, but it will enable research on syntactically annotated resources in FoLiA or psdx *xml* format.
Budget	€ 37K
Centre	Meertens
Scientific impact	The tool supports syntactic research, both synchronically and diachronically, by making it possible to query syntactically annotated corpora for syntactic construstions.
Status	The project has effectively finished and has submitted its final reports, which have been approved. [application]

MDF	Meertens Deposition Facility
Project Coordinator	Dr. Marc Kemps-Snijders (Meertens)
Goal	The project is a pilot for upgrading the Meertens deposition possibilities both in the interest of its own researchers and of the Dutch research community. It brings Meertens in a good position to host data and make that data also available for use by the research community. This has become urgent because of the reduced activities of MPI/TLA in this respect. Access restrictions as specified by the owners will be honored.
Budget	€ 40K
Centre	Meertens
Scientific impact	It will create additional options to ensure availability and accessibility of important data for the Dutch research community.
Status	The project has finished and has submitted its final reports, which have been approved. [web interface] (only open to the whole community as of the end of 2016)

OpenCGN	Incorporating CGN into OpenSONAR
Project Coordinator	Dr. Nelleke Oostdijk (Radboud University Nijmegen)
Goal	The OpenCGN project aims to provide access to the Spoken Dutch Corpus (CGN) through a web-based corpus exploitation environment, one which also gives access to the (written) SoNaR corpus.
Budget	€ 39.9K
Centre	most probably INL
Scientific impact	It will provide better empirical support for linguistic research by providing combined access to the SONAR written corpus and the Spoken Dutch Corpus (CGN).
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [application]

TPC	TaalPortaal with links to Corpora
Project Coordinator	Dr. Ton van der Wouden (Meertens Institute)
Goal	The project creates queries in corpus search interfaces and links to them from the Taalportaal. Links will be made for example sentences, grammatical terms and names or descriptions of constructions. The queries will be offered to users of Taalportaal via standard internet hyperlinks. Clicking on such a link brings the user to a corpus query interface where the specified query will be executed
Budget	€ 57.8K
Centre	Meertens
Scientific impact	It will make a direct link between grammatical descriptions and actual data illustrating the description, thus enabling testing of the claims made. It will stimulate linguists to make use of corpus data in their research and will have the educational side effect that linguists will gradually become acquainted with such corpus query applications.
Status	The project has effectively finished and has submitted its final reports, which have been approved. Several examples of query options can be found in the Taalportaal, e.g. here

1.2.6   Other Projects

Coordinator Semantic Interoperability

The work by the *Coordinator Semantic Interoperability* (Ineke Schuurman) started up in 2011 as coordinator of ISOCAT, has continued until April 1, 2015. Because of the transition within CLARIN

from ISOCAT to the CLARIN Concept Registry (CCR), we have generalized the function title. The work will be continued as of April 1st, 2015 in the CLARIAH project.

Centres of *Expertise* (WBS 1.6)

The CLARIN-NL Centre of Expertise on the curation of data, led by the Radboud University of Nijmegen but with close involvement of experts from the CLARIN Centres, discontinued the CLARIN-NL Data Curation Service already in 2014. The work has been taken up again in the context of the CLARIAH-project.

As to the second Centre of Expertise, [Nederlab](#)[2] , on data and tools for the study of the Dutch language and culture across time, the Nederlab demonstrator project finished in 2014. An initial version of the Nederlab application has been made available and was demonstrated at the CLARIN-NL Final Event on March 13, 2015.

CLARIN-TCC (CLARIN Traveling Campus Talk of Europe) is separately financed project, which has organised two Creative Camps in 2015. Its original finishing date has been extended from 31 December 2015 to May 31st 2015, and it plans to organize a workshop to come to a European research proposal for research around parliamentary data in Europe.

1.2.7   Education and Awareness (WBS 1.7)

An overview is given of the various activities for stimulating education, training, and awareness carried out in 2015.

Events visited[3]

From	To	Event	City	Ctry	Who
5-feb-15	5-feb-15	CLIN25	Antwerpen	BE	many
7-feb-15	5-feb-15	TIN-dag	Utrecht	NL	JO
24-mrt-15	24-mrt-15	Digitale Historische Kranten als Big Data	Den Haag	NL	JO, AvH
26-mrt-15	26-mrt-15	CLARIN-TCC Creative Camp	Amsterdam	NL	JO
1-apr-15	1-apr-15	Symposium *Luid zingend op een ijsschots de zomer tegemoet: de*	Amsterdam	NL	JO, HB, AvH

---

[2] http://www.nederlab.nl/

[3] JO=Jan Odijk, DB=Daan Broeder, AvH=Arjan van Hessen, HB=Hans Bennis, AD=Alice Dijkstra

From	To	Description	City	Ctry	Who
		toekomst van de geesteswetenschappen			
9-apr-15	9-apr-15	Invited Lecture at the Austrian Centre for Digital Humanities	Vienna	AT	JO
28-apr-15	28-apr-15	Digital Humanities Spring School (Keynote)	Leuven	BE	JO
20-sep-15	24-sep-15	PARSEME Workshop	Iaşi	RO	JO
25-sep-15	26-sep-15	DRONGO-festival	Utrecht	NL	JO, AvH
15-okt-15	17-okt-15	CLARIN Conference	Wrocław	PL	many
30-okt-15	30-okt-15	CLARIN-TCC Creative Camp	Den Haag	NL	JO
16-nov-15	16-nov-15	Workshop Morphosyntactische verrijking van 17e eeuwse teksten	Utrecht	NL	JO
19-nov-15	20-nov-15	CLARIN General Assembly	Copenhagen	DK	JO, AD
18-dec-15	18-dec-15	CLIN26	Amsterdam	NL	JO

At CLIN25 in Antwerp, there were lectures and posters on SHEBANQ, OpenSonar (2x), GrETEL (2x), ISOcat/ CCR, TiCCLops/Nederlab, and NameScape.
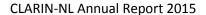
At the CLARIN Annual Meeting, which was held in Wrocław, Poland, there were about 10 presentation, posters and/or demos with authors from the Netherlands.

At CLIN26 in Amsterdam, there were about 4 lectures related to CLARIN.

Events organized by CLARIN-NL

From	To	Description	City	Ctry	Who
13-mrt-15	13-mrt-15	CLARIN-NL Final Event and CLARIAH Kick-off	Hilversum	NL	many

Publications

CLARIN-NL Book A book about CLARIN in the Low Countries (the Netherlands and Flanders) is being prepared and will be published by Ubiquity press under Open Access. Jan Odijk and Arjan van Hessen are the editors. 26 submissions for contributions to the book have been submitted, and are currently under review. Each contribution will be reviewed by at least 2 and in most cases 3 independent reviewers. The book will consist of multiple parts, provisionally characterized as follows:

- Core Infrastructure facilities
- Annotation and analysis of linguistic properties in text corpora
- Data and Tools for other linguistic research
- Data and tools for Lexical resources
- Digital data and tools other humanities disciplines

 It is expected that the book will appear by the end of 2016.

Lingua A *Lingua* special issue is being prepared, with Jan Odijk as guest editor. Seven papers have been submitted. Six of these have been accepted. Jan Odijk wrote an introduction to the Special Issue. We provide some details about the contributions:

> **Topics covered**: synchronic syntax (in 3 contributions), language variation (in 3 contributions), language acquisition, language contact, language typology, phonology/phonetics, and diachronic syntax (each in 1 contribution).[4]

> **Distribution of the primary authors by Country**: Netherlands 4; Belgium, and Germany each 1.

> **Enhanced publications**: Though Lingua explicitly invited authors to submit enhanced publications, only one contribution comes with a data set in a separate file as support for the article. It looks as if this will consist of a reference to a CLARIN Virtual Collection.

The review process has been finished. Four articles are already available on-line via the Lingua website. Two articles and the introduction are currently in production.

The articles by the non-Dutch authors are:

Hinrichs, Erhard. 2016. 'Substitute Infinitives and Oberfeld Placement of Auxiliaries in German Subordinate Clauses: A Synchronic and Diachronic Corpus Study Using the CLARIN Research Infrastructure'. To appear in *Lingua.*

Van Eynde, Frank, Liesbeth Augustinus, and Vincent Vandeghinste. 2016. 'Number agreement in copular constructions. A treebank-based investigation'. doi:10.1016/j.lingua.2016.02.001. To appear in *Lingua*.

---

[4] Many contributions cover multiple topics

Other Publications Here's a (not fully complete) overview of publications from 2015 and in part from 2016:

Barbiers, Sjef, Marjo van Koppen, Hans Bennis, Norbert Corver (2016). MIcrocomparative MOrphosyntactic REsearch (MIMORE): Mapping partial grammars of Flemish, Brabantish and Dutch. doi:10.1016/j.lingua.2015.10.018. to appear in *Lingua*.

Bouma, G., van Koppen, J. M., Landsbergen, F., Odijk, J. E. J. M., van der Wouden, T. & van de Camp, M. 'Enriching a Descriptive Grammar with Treebank Queries', In : *Proceedings of the Fourteenth International Workshop on Treebanks and Linguistic Theories (TLT14)*, 14, p. 13-25. [url] [pdf]

Brugman, Hennie, Martin Reynaert, Nicoline van der Sijs, René van Stipriaan, Erik Tjong Kim Sang and Antal van den Bosch (2016), 'Nederlab: Towards a Single Portal and Research Environment for Diachronic Dutch Text Corpora', accepted for publication in the LREC 2016 proceedings.

Cornips, Leonie, Jos Swanenberg, Wilbert Heeringa, Folkert de Vriend (2016). The relationship between first language acquisition and dialect variation: Linking resources from distinct disciplines in a CLARIN-NL project. doi:10.1016/j.lingua.2015.11.007. to appear in *Lingua*.

Heuvel, Henk van den and Nelleke Oostdijk (2016), 'Falling silent, lost for words ... Tracing personal involvement in interviews with Dutch war veterans', accepted for publication in the LREC 2016 proceedings.

Leinonen, Therese, Çağrı Çöltekin, John Nerbonne (2016). Using Gabmap. doi:10.1016/j.lingua.2015.02.004 . to appear in *Lingua*.

Odijk, J. (2015), 'Metadata Curation Strategy', manuscript Utrecht University. [pdf] with 2 appendixes: resource type [xlsx] and modality [xlsx]

Odijk, J. (2015) 'Linguistic Research with PaQU' *Computational Linguistics in The Netherlands journal* 5, p. 3-14 [pdf] [url]

Odijk, J. (2016) 'A Use Case for Linguistic Research on Dutch with CLARIN', in K. De Smedt (ed.), *Selected Papers from the CLARIN Annual Conference 2015,* 45-61. [Abstract and Fulltext]

Odijk, J. (2016), 'Linguistic Research using CLARIN – Introduction'. doi:10.1016/j.lingua.2016.04.003. To appear in *Lingua.*

Odijk, J. (2016), 'CLARIAH in the Netherlands', accepted for publication in the LREC 2016 proceedings.

Reynaert, Martin (2016), 'OCR Post-Correction Evaluation of Early Dutch Books Online – Revisited', accepted for publication in the LREC 2016 proceedings.

Schuurman, Ineke, Menzo Windhouwer, Oddrun Ohren, Daniel Zeman (2016), 'CLARIN Concept Registry: The New Semantic Registry', in K. De Smedt (ed.), *Selected Papers from the CLARIN Annual Conference 2015,* 62-70. [Abstract and Fulltext]

Sijs, Nicoline van der, Eric Sanders, Henk van den Heuvel & Aukje Borkent (2016), 'Curation of Dutch Regional Dictionaries', accepted for publication in the LREC 2016 proceedings.

Sluijs, Robbert van, Margot van den Berg, Pieter Muysken (2016). Exploring genealogical blends: The Surinamese Creole cluster and the Virgin Island Dutch Creole cluster. doi:10.1016/j.lingua.2015.12.004 . to appear in *Lingua*.

Windhouwer, Menzo , Marc Kemps-Snijders, Paul Trilsbeek, André Moreira, Bas Van der Veen and Guilherme Silva (2016). 'FLAT: constructing a CLARIN compatible home for language resources', accepted for publication in the LREC 2016 proceedings.

Wouden, Ton van der, Gosse Bouma, Matje van de Camp, Marjo van Koppen, Frank Landsbergen, Jan Odijk (2016). 'Enriching a Grammatical Database with Intelligent Links to Linguistic Resources', in K. Desmedt (ed.), *Selected Papers from the CLARIN Annual Conference 2015,* 108-117. [Abstract and Fulltext]

Wouden, Ton van der et al. Het Taalportaal: Een nieuwe wetenschappelijke grammatica voor het Nederlands en het Fries. To appear in *Nederlandse Taalkunde*, 2016

Presentations

There have been quite a number of presentations to disseminate the results obtained and experiences gained in the Netherlands. Highlights include plenary lectures at CLIN25 (Antwerp) and at TIN 2015 (Utrecht) on CLARIN, an invited lecture at the Austrian Academy of Sciences (Vienna), a key note lecture at the Digital Humanities Spring School in Leuven, a key note lecture on Multiword Expressions and LMF (CLARIN-NL DuELME project) at the PARSEME Workshop (Iaşi, Romania), and multiple presentations and demo's at the CLARIN Conference.

Here's a (most likely not fully complete) overview of the presentations given:

Boer, V. de et al. (2015), 'Dutch Ships and Sailors', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Bosch, Antal van den, Hennie Brugman en René van Stipriaan (2015), 'Lancering en demonstratie van Nederlab', March 13, 2015 Beeld en Geluid at the CLARIAH kick-off.

Bouma, G., van Koppen, J. M., Landsbergen, F., Odijk, J. E. J. M., van der Wouden, T. & van de Camp, M. 'Enriching a Descriptive Grammar with Treebank Queries', Presentation at the *Fourteenth International Workshop on Treebanks and Linguistic Theories (TLT14)*.

Bouwman, A. (2015), 'BMN-I: Linked Data on Middle Dutch Sources Kept Worldwide', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Broeder, D. (2015), 'CLAPOP, the CLARIN-NL Portal', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Camp, M. van de et al. (2015), 'OpenSONAR', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Crasborn, O. (2015) COLTIME: Presentation/demonstration at closing event of CLARIN-NL, Hilversum

Crasborn, O. (2015) EXILSEA: Presentation/demonstration at closing event of CLARIN-NL, Hilversum.

Dijk, S. van (2015), 'COBWWWEB', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Goosen, Twan, Thomas Eckart, Matej Durco, Davor Ostojic (2015), 'VLO developments', presentation and demo at the CLARIN Conference, Wrocław, Poland, 16 October 2015.

Gorp, J. van (2015), 'AVResearcherXL', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Komen, Erwin (2016).   Corpus Studio Web Version. Demo and poster at the CLARIAH-dag, Amersfoort, January 22.

Landsbergen , Frank, Ton van der Wouden, Jorik van Engeland, Erlinde Meertens, en Sylvia Wijsman (2015), 'Enriching a grammatical database with intelligent links to linguistic resources', demo at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015.

Mak, I. et al (2015), 'CORNETTO', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Niestad, J. et al. (2015), 'AutoSearch', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Niestadt, Jan (2015), BlackLab: advantages and disadvantages of using Lucene for a linguistic corpus search system. Presentation at CLIN26, Amsterdam, 18 December 2015.

Noord, G. van (2015), 'PaQu', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Noord, Gertjan van & Jan Odijk (2016). `Goed of Fout', demonstration at the CLARIAH-day, Amersfoort, January 22, 2016. [pdf] [pptx]

Noord, Gertjan van & Jan Odijk (2016). `Goed of Fout: Wat gebruikt men feitelijk?', presentation at the  `Grote Taaldag' (TIN), Utrecht, 6 February 2016. [handout] [pptx] [pdf]

Odijk, J. (2015), 'CLARIN-NL: Contributions by NLP & Current Status', presentation at CLIN25 2015, Antwerp, Feb 6, 2015 [pptx] [pdf]

Odijk, J (2015), 'Linguistic Research with PaQu', poster at CLIN25 2015, Antwerp, Feb 6, 2015 [pptx] [pdf]

Odijk, J. (2015), 'What CLARIN has to offer to linguists', plenary lecture at the TIN-dag, Utrecht, Feb 7, 2015. [pptx] [pdf]

Odijk, J. (2015), 'Investigating *heel, erg, zeer* with PaQu', lecture at the TIN-dag, Utrecht, Feb 7, 2015. [pdf]

Odijk, J. (2015), 'CLARIAH: A Research Infrastructure for the Humanities in the Netherlands', invited lecture at the Austrian Academy of Sciences, Vienna, 9 April 2015. [pptx]

Odijk, J. (2015), `Infrastructure for Digital Humanities', lecture at the Utrecht Digital Humanities Session, Utrecht, 10 June 2015. [pptx]

Odijk, J. (2015), 'Multiword Expressions and LMF', invited lecture held at the PARSEME Hands-on workshop on the lexical encoding of multiword expressions, Iaşi, Romania, 22 September 2015. presentation: [ppt] [pdf] ; handout: [pdf]

Odijk, J. (2015), 'PARSEME Alpino MWE Encoding', lecture held at the PARSEME Hands-on workshop on the lexical encoding of multiword expressions, Iaşi, Romania, 23 September 2015. [ppt] and dataset

Odijk, J. (2015), 'Zoeken naar Constructies', presentation and poster held at the DRONGO Language Festival, Utrecht, 26 September 2015. [presentation] [poster]

Odijk, J. (2015), 'Taal- en Spraaktechnologie (TST) voor het Nederlands: Status 2015', presentation held at the Taalcongres 'Taal schept Kansen', Flemish Parliament, Brussels, 10 October 2015. [pptx]

Odijk, J. (2015), 'Using PaQu for language acquisition research', presentation at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015 [pptx] [pdf]

Odijk, J. (2015), 'Using PaQu for language acquisition research', demo at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015 [pdf]

Odijk, J (2015), 'Can we extend Alpino with flexible MWEs? An unexpected possible use of PaQu', presentation and demo at the CLARIN Conference, Wrocław, Poland, 16 October 2015. [pdf] [data] [handout] [analysis]

Odijk, J. (2015), 'Linguistic Research with CLARIN', MA Rotation Lecture, Utrecht, 2015-11-10. [pptx] [pdf]

Odijk, J. (2015), 'Extending Alpino with Flexible MWEs. An unexpected possible use of PaQu', poster and presentation/demo at CLIN26, Amsterdam, 18 December 2015 [poster] [presentation]

Peursen, W. van & D. Roorda (2015), 'SHEBANQ', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Peursen, Wido van, Dirk Roorda and Martijn Naaijer (2015), 'SHEBANQ: Annotations based hub of research in the Hebrew Bible'. Presentation at CLIN26, Amsterdam, 18 December 2015.

Reynaert, M. et al. (2015), 'PICCL: Philosophical Integrator of Computational and Corpus Libraries'', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Reynaert, Martin, Maarten Van Gompel, Ko van der Sloot and Antal van Den Bosch (2015), 'PICCL: Philosophical Integrator of Computational and Corpus Libraries'. Presentation at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015.

Reynaert, Martin & Matje van de Camp (2016), OpenSoNaR-CGN. Poster and demo at the CLARIAH-dag, Amersfoort, 22 January.

Roorda, Dirk (2015), 'Hebrew Bible: System of Data and Annotations', presentation and demo at the CLARIN Conference, Wrocław, Poland, 16 October 2015.

Schuurman, I, & M. Windhouwer (2015), 'The CLARIN Concept Registry (CCR)', poster and demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Schuurman, Ineke, Menzo Windhouwer, Oddrun Ohren & Daniel Zeman (2015). 'CLARIN Concept Registry: the new semantic registry', presentation at the CLARIN 2015 Conference, Wrocław, Poland, 17 October 2015. [pptx] [pdf]

Schuurman, Ineke, Menzo Windhouwer and Olha Shkaravska (2015). 'The CLARIN Concept Registry'. Presentation at CLIN26, Amsterdam, 18 December 2015.

Sijs, Nicoline van der, Henk van den Heuvel, Roeland van Hout & Eric Sanders (2016), 'CARE: Curation of Dutch Regional Dialect Dictionaries', demo en poster presentation at the CLARIAH-dag , Amersfoort, 22 January.

Sloetjes, H., Seibert, O., Somasundaram, A., & Drude, S. (2015). Language identification in ELAN. E-poster presented at the 4th International Conference on Language Documentation & Conservation (ICLDC), Honolulu, HI.

Verberne, S. (2015), 'RemBench', demo at the CLARIN-NL Closing Event, Hilversum, March 13, 2015.

Windhouwer, Menzo (2015), 'CMDI first aid kit', presentation and demo at the CLARIN Conference, Wrocław, Poland, 16 October 2015.

Wouden, Ton van der, Gosse Bouma, Matje van de Kamp, Marjo van Koppen, Frank Landsbergen and Jan Odijk (2015), 'Enriching a grammatical database with intelligent links to linguistic resources', presentation at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015 [pdf]

Wouden, Ton van der and Frank Landsbergen (2016). ''Enriching a grammatical database with intelligent links to linguistic resources'. Poster and demo at the CLARIAH-dag, Amersfoort, 22 January.

Wouden, Ton van der (2016). A boost for the digital infrastructure for linguistics: the Taalportaal. DH Benelux 2016.

Wouden, Ton van der (2016). Taalportaal: A new tool for linguistic research. Submitted & accepted DH 2016 Krakow.

Wynne, Martin and Arjan van Hessen (2015), 'Working with oral history archive data', poster at the CLARIN 2015 Conference, Wrocław, Poland, 16 October 2015.

Support for other events

From	To	Event	City	Form of Support
5-feb-15	5-feb-15	CLIN25	Antwerpen	financial
13-apr-15	13-apr-15	Workshop STCN	Den Haag	financial
25-sep-15	26-sep-15	DRONGO Festival	Utrecht	financial (via UU)
16-nov-15	16-nov-15	Workshop Morphosyntactische verrijking van 17e eeuwse teksten	Utrecht	Website, announcements, active participation
18-dec-15	18-dec-15	CLIN26	Amsterdam	financial

Support for travel of Dutch researchers to CLARIN-related events
None in 2015.

CLARIN- NL Newsflashes

CLARIN- NL Newsflashes have been sent out to all registered persons on the following dates:

- NewsFlash 32 (12 jan 2015)
- NewsFlash 33 (13 feb 2015)
- NewsFlash 34 (6 mrt 2015)
- NewsFlash 35 (9 sep 2015)
- NewsFlash 36 (21 okt 2015)

CLARIN-EU Newsflashes

In 2015, no contributions to the CLARIN (ERIC) newsflashes have been made.

CLARIN-EU Newsletter

In 2015 no CLARIN (ERIC) Newsletters have appeared.

eData & Research

CLARIN-NL continued its participation in the eData & Research Magazine in 2015. Approximately 6 articles on CLARIN-NL or related topics have appeared in eData&Research in 2015. Erica Renckens participates in the editorial team and Arjan van Hessen in the board.

The articles that have been published in eData & Research in 2015 can be accessed via http://www.clarin.nl/node/188.

A dual interview with Jan Odijk and Jaap van de Herik on the occasion of the end of CLARIN-NL and CATCH projects will appear in the eData & Research Magazine in June 2016.

Movies and Screen Captures

Several projects have made movies or screen captures to illustrate their work or their demonstrators. Rembench also added such a short movie on its help page:

Project	Link to Movie / Screen capture
GabMap	http://www.youtube.com/watch?v=anElatt5fFM&feature=plcp
Interviews / Veteran Tapes	http://www.youtube.com/watch?v=VpCXh_NVVso
WFT-GTB	http://www.youtube.com/watch?v=fevAPknCaDo&feature=player_embedded
RemBench	http://rembench.huygens.knaw.nl/help

OpenSONAR and @PhilosTEI each have a guided tour through the application via their Help button

The EB commissioned the production of 9 short movies on CLARIN-NL projects (to Dyzlo). Below one can find the CLARIN-NL projects that have been selected for these movies. All movies are available.

All movies and screen captures can be reached via http://www.clarin.nl/node/403

Project	Link to Movie
MIMORE	http://vimeo.com/35305355 http://www.youtube.com/watch?v=TBDGrqKYbIA
SignLinc	http://vimeo.com/35305490 http://www.youtube.com/watch?v=je3ZKPw6w5s
Arthurian Fiction	http://vimeo.com/35305561 http://www.youtube.com/watch?v=t0y-ecrVD5w
WFT-GTB	http://vimeo.com/39044483 http://www.youtube.com/watch?v=aSWKJucbWY8
WIP	http://vimeo.com/48195599 http://www.youtube.com/watch?v=tEqAIH2o2rc&feature=plcp
Search & Develop	http://vimeo.com/40289489 http://www.youtube.com/watch?v=6CkCfaJ7rao&feature=plcp
CKCC	http://vimeo.com/68203559 http://www.youtube.com/watch?v=Sl0Ig7SCOcc
PILNAR	http://vimeo.com/100788014 https://www.youtube.com/watch?v=iMWhO3D_n_Q
Polimedia	http://vimeo.com/100787499 https://www.youtube.com/watch?v=u24oRCj7xrQ

Three new movies were planned for 2015. These are currently being finalized.

They are movies about the projects D-LUCEA, DSS, and OpenSONAR. A preliminary version of the D-LUCEA movie (not subtitled yet) is accessible via Vimeo: https://vimeo.com/152447934 (password required), but all movies will soon be accessible via the CLARIN-NL portal.

The movies have been made available in streaming mode via Vimeo and YouTube, have subtitles (at least in English) and the originals have been made available on solid state memory.

Website
The website was maintained by an external company (*Civinex*) until the end of 2015.
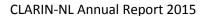
The CLARIN-NL website is a project-internal website. We will try to keep it live for as long as possible (but have stopped the maintenance). For aspects that go beyond the project, we will use the CLARIN-NL portal, a preliminary version of which has gone live in June 2014 and which has been officially launched at the CLARIN-NL final event.

Education
Various educational activities have been organized in 2014. The following table lists the most important ones:

From	To	Event	City	Ctry	Who
Dec 2014	Jan 2015	Corpus Analysis Course	Utrecht	NL	JO
12-jan-15	16-jan-15	LOT Winterschool Course: linguistic research Using CLARIN	Amsterdam	NL	JO and Sjef Barbiers
19-mrt-15	19-mrt-15	*Leren werken met een nieuwe zoektool voor de Short-Title Catalogue Netherlands*	Den Haag	NL	Els Stronks
13-okt-15	13-okt-15	Gastcollege DH Minor: CLARIN	Utrecht	NL	JO
10-nov-15	10-nov-15	MA-Rotation Lecture: CLARIN	Utrecht	NL	JO
Nov 2015	Jan 2016	Corpus Analysis Course	Utrecht	NL	JO

Helpdesk

The CLARIN-NL Helpdesk[5] has run all year. The amount of requests is very low, and therefore the Helpdesk was easily managed in 2015 by Jan Odijk and Daan Broeder. There are more frequent informal requests directly to these persons from researchers who inquire about CLARIN and its use.

1.2.8 National Coordination Point

The national coordination point has functioned well in 2015. Prof.dr. Jan Odijk is the Programme Director. Drs. Arwin van der Zwan was the project secretary.

The governance structure consists of an Executive Board (EB), Board, National Advisory Panel (NAP) and International Advisory Panel (IAP). Their tasks and responsibilities have been laid down in the consortium agreement.

Executive Board The Executive Board consists of 4 members:

* Prof.dr. Jan Odijk, programme director

* Prof.dr. Hans Bennis, chair and representative of the humanities researchers

* Dr. Arjan van Hessen, responsible for complementary activities such as PR, education, training and awareness creation.

* Ir. Daan Broeder, technical director

The executive board has held seven meetings in 2015.

Board The Board consists of 8 members. A full list of its members can be found in Appendix A.

The current composition of the Board can also be found on the CLARIN-NL website (http://www.clarin.nl/category/board/board)

The Board has held no meeting in 2015. The final meeting of the CLARIN-NL Board is scheduled for May 9, 2016.

NAP The National Advisory Panel consists of 16 members, providing a good representation of the fields of the humanities, language and speech processing and infrastructure design and implementation. A full list of the members can be found in Appendix B

The current composition of the CLARIN-NL NAP can also be found on the CLARIN-NL website (http://www.clarin.nl/category/board/national-advisory-panel).

The NAP has held its final meeting in 2015 (Feb 17).

IAP The International Advisory Panel consists of 11 members. A full list of the members can be found in Appendix C and on the CLARIN-NL website (http://www.clarin.nl/category/board/international-advisory-panel).

The IAP has held no meetings in 2015 but was invited for the CLARIN-NL Final event and the CLARIAH Kick-off (March 13).

---

[5] http://trac.clarin.nl/trac

CLARIN-NL Long Term Working Plan The CLARIN-NL Long Term Working Plan was approved by the Board on August 19, 2009, and has remained unchanged. It can be found on the CLARIN-NL website: http://www.clarin.nl/system/files/CLARIN-NL%20Multiyear%20Programme%20090409-2.pdf

Working Plan 2015 The Working Plan for 2015 was approved by the Board via e-mail. It can be found on the CLARIN-NL website (http://www.clarin.nl/page/about/47)

Overall, the governance structure as set up functions appropriately.

4   2014 Finances

A detailed report on the CLARIN-NL finances in 2015 can be found in the 2015 Financial Report.
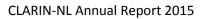
5   Other Activities

As a response to the invitation of the *Commissie voor de update van de Nationale Roadmap voor Grootschalige Onderzoeksfaciliteiten* (Committee for the update of the National Roadmap for Large Scale Research Facilities), CLARIN-NL and the Dutch players in DARIAH have joined forces and formulated a new proposal for a large scale national infrastructural project called *Common Lab Research Infrastructure for the Arts and the Humanities (CLARIAH)*, uniting all humanities organizations in the Netherlands. Details about it can be found on the CLARIAH website (www.clariah.nl). The new project, called CLARIAH-CORE, has started on January 1, 2015 and will run for four years.

6   Research and Education

It is clear that the CLARIN infrastructure is increasingly often used in actual research and in education. Research that crucially used the CLARIN infrastructure is now appearing in the Lingua Special Issue on the CLARIN Infrastructure with Jan Odijk as guest editor. We are approached by many researchers when they are preparing new project proposals. The research project proposal by Van Koppen and Dietz for the NWO *Vrije Competitie* programme, in which we contributed to the preparation and in which some of us participate actively   was awarded funding. CLARIN (JO) is involved in a project proposal for the NWO *Gravitation* programme that is being prepared by a large consortium of linguists united in LOT, and, if it is awarded, the researchers will make heavy use of the CLARIN Infrastructure and   also contribute to it. The first PhD in which crucial use is made of CLARIN has been successfully defended by Liesbeth Augustinus in Leuven (18 October 2015). Though this is a PhD in Flanders, the work on CLARIN in Flanders was made possible by the close collaboration of Flanders with CLARIN-NL.

There have been many educational activities, including a LOT Winterschool course, and multiple lectures on specific aspects of CLARIN at least in Utrecht and in Amsterdam (UvA). Since some of these are part of the regular curriculum, they will continue also now that the CLARIN-NL project has finished. Others will be continued through the CLARIAH-CORE project.
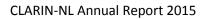
Appendix A Composition of the Board

Name	Affiliation
Prof.dr. Geert Booij (until Sep 1, 2012)	LUCL, Universiteit van Leiden
Prof.dr. Lou Boves	CLST, Radboud Universiteit Nijmegen
Dr. Peter Doorn	DANS
Prof.dr. Martin Everaert	UIL-OTS, Universiteit van Utrecht
Prof.dr. Jaap van den Herik	TiCC, Universiteit van Tilburg
Prof.dr. Aafke Hulk (until Sep 1, 2012)	ACLC, Universiteit van Amsterdam
Prof.dr.ir John Nerbonne	CLCG, Universiteit van Groningen
Prof.dr. Els Stronks (as of Sep 1, 2012)	Universiteit van Utrecht
Ir. Peter Wittenburg (as of Sep 1, 2012)	Max Planck Instituut voor Psycholinguïstiek, Nijmegen

Appendix B Composition of the NAP
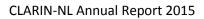
Name	Affiliation
Prof.dr. Willem Adelaar (until Sep 1, 2012)	LUCL, Universiteit van Leiden
Prof.dr. Sjef Barbiers	Meertens Instituut
Dr. Jeannine Beeken	INL
Prof.dr Antal van den Bosch	Radboud Universiteit Nijmegen
Dr. Gosse Bouma	Universiteit van Groningen
Drs. Hennie Brugman	Meertens Instituut
Prof.dr. Karina van Dalen-Oskam	Huygens Instituut voor Nederlandse Geschiedenis
Dr. Paul Doorenbosch	Koninklijke Bibliotheek
Dr. Willemijn Heeren	Universiteit van Twente
Ir. Marc Kemps-Snijders	Meertens Instituut
Prof.dr. Kees Mandemakers	Internationaal Instituut voor Sociale Geschiedenis
Dr. Nelleke Oostdijk	CLST, Radboud Universiteit Nijmegen
Prof.dr. Reinier Salverda	Fryske Akademy
Prof.dr. Ted Sanders (until Sep 1, 2012)	UiL-OTS, Universiteit Utrecht
Prof.dr. Piek Vossen	Vrije Universiteit Amsterdam
Prof.dr. Sally Wyatt (until Sep 1, 2012)	eHumanities (KNAW)
Dr. Joris van Zundert	Huygens Instituut voor Nederlandse Geschiedenis
Prof.dr. Arjen Versloot (as of Sep 1, 2012)	Universiteit van Amsterdam
Prof.dr. Toine Pieters (as of Sep 1, 2012)	Universiteit van Utrecht

Appendix C Composition of the IAP

Name	Affiliation	Country
Prof.dr Scott Farrar	University of Washington, Department of Linguistics	USA
Prof.dr David Hoover	New York University	USA
Prof.dr. Clifford Lynch	Coalition for Networked Information	USA
Prof.dr. Harold Short	King's College London, Centre for Computing in the Humanities	UK
Prof.dr. Benjamin T'sou	City University of Hong Kong	China
Prof.dr Hugo Van hamme	Katholieke Universiteit Leuven	Belgium
Prof.dr. Tamás Váradi	Linguistics Institute, Hungarian Academy of Sciences	Hungary
Prof.dr. Koenraad De Smedt	University of Bergen	Norway
Prof.dr. Erhard Hinrichs	University of Tübingen	Germany
Prof.dr. Jan Hajič	Charles University Prague	Czech republic