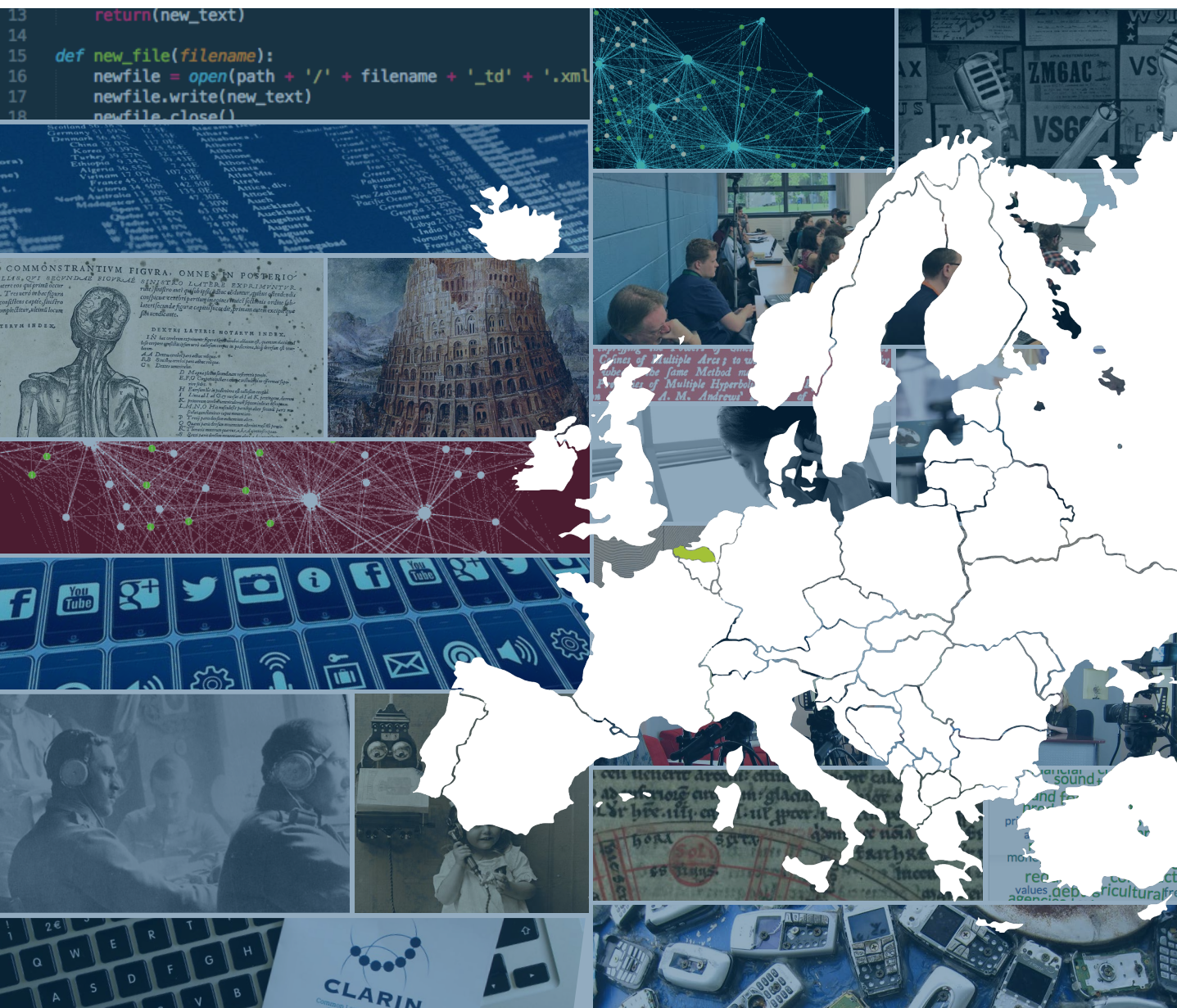


# Tour de CLARIN

## DLU / Flanders



Written by Catia Cucchiarini, Griet Depoorter, Katrien Depuydt, Ineke Schuurman, Leen Sevens, Hans Westgeest, Darja Fišer and Jakob Lenardič, and edited by Darja Fišer and Jakob Lenardič

# Foreword

Tour de CLARIN highlights prominent user involvement activities of CLARIN national consortia with the aim to increase the visibility of CLARIN consortia, reveal the richness of the CLARIN landscape, and display the full range of activities throughout the CLARIN network that can inform and inspire other consortia as well as show what CLARIN has to offer to researchers, teachers, students, professionals and the general public interested in using and processing language data in various forms.

This brochure presents DLU / Flanders and is organized in five sections:

- Section One presents the members of the consortium and their work
- Section Two demonstrates an outstanding tool
- Section Three highlights a prominent resource
- Section Four reports on a successful event for researchers and students
- Section Five includes an interview with a renowned researcher from the digital humanities or social sciences who has successfully used the consortium's infrastructure in their research



Ghent, Belgium | photo by Jochen Schaft | Pixabay



# DLU/Flanders

*Written by Catia Cucchiarini, Ineke Schuurman and Griet Depoorter, and edited by Darja Fišer and Jakob Lenardič*

CLARIN DLU/Flanders<sup>1</sup> is a founding member of CLARIN and represents Flanders, the Dutch-speaking part of Belgium. The consortium consists of

- the Dutch Language Union (Nederlandse Taalunie - DLU);
- the Dutch Language Institute (Instituut voor de Nederlandse Taal – INT);
- the Centre for Computational Linguistics (Centrum voor Computerlinguïstiek – CCL – University of Leuven);
- the Language and Translation Technology Team (LT<sub>3</sub> - University of Ghent);
- the Computational Linguistics & Psycholinguistics (CLiPS) research group (University of Antwerp);
- the PSI Speech Group (ESAT-PSI – University of Leuven); and
- the Language Intelligence & Information Retrieval research lab (LIIR – University of Leuven).

Since Flanders is not a country but a region, it did not qualify as a member in CLARIN and is therefore represented by the Dutch Language Union (DLU), an international language policy organisation. The consortium is coordinated by the Dutch Language Institute (INT) and the national coordinator Griet Depoorter. The INT is a starting point for anyone who wants to know anything about the Dutch and Flemish languages through the centuries. The institute takes a central position in the Dutch-speaking world as a developer, manager and distributor of sustainable language resources, using reliable scholarly methods, and is a certified CLARIN B-Centre. For instance, the INT produced the *Woordenboek der Nederlandsche Taal* (Dictionary of the Dutch language), an enormous historical

dictionary which describes Dutch words from 1500 to 1976. A few other examples of resources and tools that are available at the INT:

- the Dictionary of Contemporary Dutch (Algemeen Nederlands Woordenboek), which is online and corpus-based;
- the Reference Lexicon Dutch (Referentiebestand Nederlands), which contains 50,000 frequent Dutch words, enriched with linguistic information;
- the Dutch Parallel Corpus, a high quality sentence-aligned parallel corpus of 10 million words for the language pairs Dutch-English and Dutch-French;
- the Word list of the Dutch Language (Woordenlijst Nederlands), a list of words in the correct official spelling; and
- Blacklab, an open source corpus search engine built on top of Apache Lucene.

Apart from providing state-of-the-art language resources and tools, the consortium is also active in involving both students and researchers in its activities. For instance, a workshop was held at the Dutch Language Institute in October 2017, aimed at familiarising digital humanities researchers with the resources and tools the consortium offers. You can read more about the workshop on page 9.



Griet Depoorter, National Coordinator of CLARIN DLU/Flanders, and Vincent Vandeghinste, National User Involvement Representative.

<sup>1</sup> <http://www.ivdnt.org/>

## Text2Picto and Picto2Text tools

*Written by Griet Depoorter, Darja Fišer, Jakob Lenardič, Ineke Schuurman and Leen Sevens*

Text2Picto and Picto2Text<sup>2</sup> are two complementary translation tools aimed at enhancing communication for people with reading disabilities. Both tools have been developed by the Centre for Computational Linguistics at the University of Leuven. Text2Picto translates sentences into pictographs – that is, graphic symbols that serve as stand-ins for verbal communication – while Picto2Text does the reverse by allowing users to select the pictographs that they want to translate into written text. Two different sets of pictographs are available for both tools – the Beta set and Sclera set. The symbols in each set are designed to be very concrete and easy to interpret, so their use reduces the cognitive complexity of reading e-mails, web pages, chats and work documents. Currently, the tools are available for Dutch, Spanish and English, but other languages can be used as well if there is a Wordnet available for them. Figure 1 shows the use of Text2Pico based on the Beta set and Figure 2 on the Sclera set.

Work is now being done by Leen Sevens from the University of Leuven to add additional features to the tools. Demos of these features are already available for Dutch:

- Spelling Correction, which is a crucially important feature since the users of Text2Picto often spell phonetically;
- Word-Sense Disambiguation, which identifies the correct sense of polysemous words and retrieves the correct pictograph for that sense; and
- Text2Picto + Simplification and Temporal Detection for Dutch (2017), which adds pictographs depicting temporal relations between the other symbols and simplifies syntactic structure (Figure 3).

The tools have generated a lot of interest and are already being used by their target audience. They were awarded the prestigious Language Industry Award in 2016 and have been implemented into the WAI-NOT website,<sup>3</sup> which allows people with mental disabilities to use the Internet within an accessible online environment. By using the pictographs, they can play games and chat even if they aren't able to read. There is a YouTube video available in Dutch that demonstrates the implementation in the WAI-NOT website. Additionally, the software is also part of the ABLE social services app, which is available through Google Play.

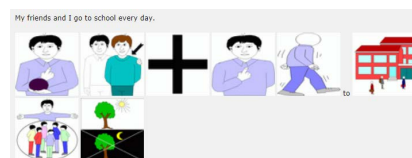


Figure 1: Translation of the sentence “My friends and I go to school every day” into the Beta pictograph set.



Figure 2: Translation of the same sentence into the Sclera set.



Figure 3: Translation of the Dutch sentence “Ik zal de rode wijn drinken die mijn moeder gekocht heeft” (“I will drink the wine that my mother has bought”) into the Sclera set. Note the second pictograph in the upper row indicating future time reference corresponding to “zal” (“will”), and the middle pictograph in the lower row indicating the past time reference of buying the wine.

<sup>2</sup> <http://picto.ccl.kuleuven.be/>

<sup>3</sup> <https://www.wai-not.be/>

## The Corpus of Contemporary Dutch

*Written by Griet Depoorter, Katrien Depuydt and Hans Westgeest, edited by Darja Fišer and Jakob Lenardič*

The Corpus of Contemporary Dutch (Corpus Hedendaags Nederlands - CHN)<sup>4</sup> is a collection of more than 800,000 texts taken from various sources, such as newspapers, magazines, news broadcasts, legal writings, and books, for the period between 1814 and 2013.

Since 1994, the Institute for Dutch Lexicology (which transformed itself into the Dutch Language Institute) made several corpora of contemporary Dutch available online: the 5, 27 and 38 Million Word Corpora and the Dutch Parole Corpus 2004. These older corpora were merged and a considerable amount of more recent material was added from the NRC Handelsblad, which is a Dutch newspaper, and De Standaard, which is a Flemish newspaper. Other sources that were added came from Suriname and the Netherlands Antilles (where Dutch is also an official language), such as newspapers, material published on internet (blog, website) and books written by Surinam authors. This collection of data became the Corpus of Contemporary Dutch, which serves as the first step towards a monitor corpus for contemporary Dutch.

The corpus contains approximately 440 million tokens:

- 224 million Dutch Dutch;
- 185 million Belgian Dutch (Flemish);
- 14.4 million Dutch as spoken in the Antilles; and
- 18.3 million Surinamese Dutch.

The corpus has been lemmatised and PoS-tagged. The CHN can be searched via a simple search interface and via CQL, and users can search for or filter by five criteria: title, author, year of publication, medium and language variety. The possible values for the last criterion are NN (Dutch from the Netherlands), BN (Dutch from Belgium), SN (Dutch from Suriname) and AN (Dutch from Netherlands Antilles). The software powering the CHN website was developed at the Dutch Language Institute during the course of the IMPACT and CLARIN projects, and the corpus search is powered by BlackLab. Another great advantage of the Corpus of Contemporary Dutch is that it continues to grow, and during 2018 a significant amount of new data will be added (among others, newspaper data from the period between 2014 and 2017).



Figure 4: Searching for all word forms of the lemma “gezellig” (“cosy”) in Flemish newspapers from 2000 to 2010.

<sup>4</sup> <http://chn.inl.nl/>

The corpus data have already been successfully used in linguistic research. Jaspers et al. (2015) used CHN data when researching the syntactic and semantic characteristics of Dutch scalar modifiers denoting small degrees (like few in English), while Devos (2016) used the corpus to investigate a special category of Dutch infinitival phrases that act as obligatory modifiers of nominal predicates and add a causative meaning to the clause. In addition, the corpus has also served as the main source of data for a number of students’ theses. As examples of using the corpus in student work, Saskia Lubrun at the University of Leiden researched the collocational properties of the Dutch subjunctive, while Wanda Polak at the University of Amsterdam investigated the phonological contexts of several Dutch suffixes.

References:

Devos, F. (2016). ‘Joost was het gewoon om dergelijke zinnen te analyseren’. Over beknopte bijzinnen als oorzakelijk object in het Nederlands. In Een Sextant Voor Een Taalspecialist: Bijdragen Tot Joost Buysschaert in Profiel, 39–44.

Jaspers, D., van Craenenbroeck, J., and Vanden Wyngaerd, G. (2015). De syntaxis en semantiek van diminishers in het Nederlands en het Duits. [https://static1.squarespace.com/static/5217e223e4b090faa01f8f2d/t/5637cab6e4b0e17c588d8538/1446496950714/paper\\_ewa\\_finaal.pdf](https://static1.squarespace.com/static/5217e223e4b090faa01f8f2d/t/5637cab6e4b0e17c588d8538/1446496950714/paper_ewa_finaal.pdf)

Lubrun, S. (2015). De collocaties van de Nederlandse subjunctief. <https://openaccess.leidenuniv.nl/handle/1887/35053>

Polak, W. (2015). In welke fonologische context komt afleiding met de achtervoegsels -ig, -erig en -achtig voor? <http://www.fon.hum.uva.nl/archive/2015/2015-BA-WandaPolak.pdf>

Query: lemma="gezellig" - Duration: 50ms

Per Hit Per Document Hits grouped Documents grouped Total hits: 3855 Total pages: 74

Prev 1 2 3 4 5 6 7 8 9 10 11 ... Next Show/hide like

Left context	Hit text	Right context	Lemma	Part of speech
Stage beauty (2004)	... op zoek is naar een <b>gezellig</b>	avondje uit zal zich aan ...	gezellig	AA(degree=pos, position=prenom)
Spookrijders op een autofoze zondag (2004)	... intellectuelen laten wekken uit hun <b>gezellige</b>	coma? Wat voor zin heeft ...	gezellig	AA(degree=pos, position=prenom, formal=infl-e)
Grote terrassen in Burgemeester Reynaertstraat (2004)	... krijgt de straat een veel <b>gezelliger</b>	aanblik. Het zal bij mooi ...	gezellig	AA(degree=comp, position=prenom)
Grote Markt klaar voor zomer (2004)	... omdat die de stad een <b>gezellige</b>	uitstraling geven", zegt Neefs. „De ...	gezellig	AA(degree=pos, position=prenom, formal=infl-e)
Er moet toch iemand mij zijn' (2004)	... om levenloze objecten, een tv-scherm, <b>gezellige</b>	oude boeken en zo." Eenzaamheid ...	gezellig	AA(degree=pos, position=prenom, formal=infl-e)
Kreunen onder een gevoel van onrecht (2004)	... weet zitten de erfgenamen straks <b>gezellig</b>	samen in de regering. Kunnen ...	gezellig	AA(degree=pos, position=adv(pred)
Veel woorden en een relatie (2004)	... zich aan een danspasje. Heel <b>gezellig</b>	allemaal, tot een man in ...	gezellig	AA(degree=pos, position=adv(pred)
Klokken en boeken (2004)	... omgetoverd tot een frisse maar <b>gezellige</b>	leeshoek. Het mooiste moment was ...	gezellig	AA(degree=pos, position=prenom, formal=infl-e)

Figure 5: Some occurrences of the lemma “gezellig”.

Workshops for Introducing Digital Humanities Researchers to the CLARIN Services & Resources

Written by Griet Depoorter and Katrien Depuydt, edited by Darja Fišer and Jakob Lenardič

In 2016, the Institute for Dutch Language (INT) sent out an information sheet and a questionnaire to Flemish research groups in the field of humanities. The goal was to promote CLARIN and the INT (as a CLARIN Centre), to get an overview of the usage of the CLARIN infrastructure and services, and to receive input concerning the expectations of the user community about CLARIN and the INT. After receiving the feedback, the director of the INT, Frieda Steurs, conducted talks with a number of these groups (for instance, the Language Group Flemish Sign Language at the University of Leuven, the Ghent Centre for Digital Humanities at Ghent University and the Department of Linguistics at the University of Antwerp) to receive further in-depth information. What the centre learnt this way is that the linguistic community in Flanders would like the CLARIN consortium to expand existing datasets (e.g. the Dutch Parallel Corpus) or create new ones, like a corpus of spoken (or written) Dutch and videos of the corresponding Flemish sign language. In addition, the INT CLARIN Centre should serve as a knowledge centre for standards and annotation protocols, as well as offer expertise and support for researchers who have little or no experience with digital research.

As a result of this knowledge sharing and gathering initiative, the INT organised the first training workshop<sup>5</sup> in October 2017 in Antwerp in cooperation with Digital Humanities Flanders (DHuF) and the Faculty of Arts of the University of Antwerp. An invitation was sent out to all DHuF members, and the event announcement was also published on the websites of DHuF and the INT. Eventually 32 researchers, primarily historians, from Antwerp, Leuven, Ghent and the Netherlands attended the workshop.

The workshop showcased what the INT, which had then been just established, can offer digital humanities researchers, especially those researchers who use historical language material. The event started with a presentation of CLARIN in general and the INT as a CLARIN centre, and continued with an introduction into corpus building, focusing on historical corpus building in particular. The best practices with regards to building historical resources were discussed and exemplified with a concrete use case; namely, with the Nederlab text collection. Nederlab is a web environment for researchers and students who study the evolution of the Dutch language, literature and culture. The website offers millions of pages of (historical) Dutch texts that can be researched and analysed with user-friendly text analysis software.

Another topic of the workshop was the enrichment of corpus material. The specific challenges of annotating historical texts were discussed; for instance, the fact that standard (modern) tag sets cannot be applied to historical texts, and that tokenisers cannot handle clitics well. Furthermore, the webservice INL Labs was demonstrated, which linguistically annotates (historical) texts (in various input formats). INL Labs uses two annotation tools: the Stanford NE tagger and the INT-developed tagger-lemmatiser for historical Dutch.

The workshop also demonstrated how historical data can be searched with Blacklab, a corpus retrieval engine which is available as a webservice and as a Java library. The tool allows fast, complex searches with accurate hit highlighting on large, tagged and annotated bodies of text and it can be used to search historical corpora like the Corpus Gysseling and Letters as Loot. Users were also shown how to search their own data by means of Autosearch, which is powered by the Blacklab Engine.

Finally, there was a presentation on the work on the diachronic computational lexica of the INT, GiGaNT and DiaMaNT, and what the benefits would be of having the data available as Linked Open Data. This was the first of several workshops that the INT CLARIN Centre plans to give in Flanders. On March 20, the consortium organised an information session that involved humanities researchers from the University of Leuven. During the session, the consortium gathered information on what the specific needs of the attending researchers are in relation to the services provided by the INT and CLARIN in general. Based on their feedback, the DLU then plans to organise follow-up workshops that will be specifically tailored to the needs of focused research groups with shared interests.

<sup>5</sup> <http://uahost.uantwerpen.be/platformdh/index.php/event/int-workshop-antwerp/>



## Cora Pots

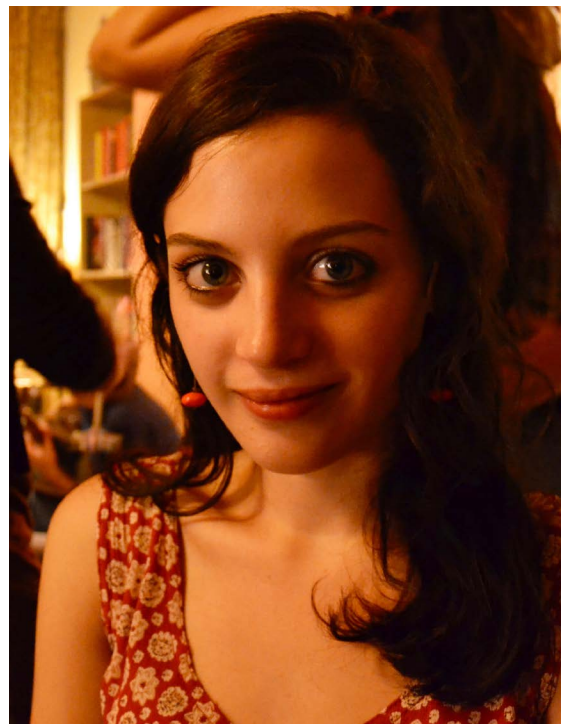
***Cora Pots is a PhD student in the Quantity and Quality in Linguistics project at the University of Leuven. The following interview took place via Skype and was conducted on 19 March 2018 and transcribed by Jakob Lenardič, edited by Darja Fišer.***

### 1. Could you briefly describe your academic background and your current position?

My interest in linguistics started when I was studying for my Bachelor's Degree at Utrecht University. It was a two-track programme in modern literature and linguistics. In the linguistics track, I was working in the generative framework, primarily child language acquisition and syntax. I did an internship where I researched speech perception and language development in younger children, which really sparked my interest in research and inspired me to pursue it as a career choice. In my Master's degree studies I slightly changed the focus of my research and began investigating syntactic variation in Germanic languages. When I became a Research Assistant at the Meertens Institute, under the supervision of Professor Sjef Barbiers, I began combining my formal background with a computational approach to linguistics and co-wrote the Educational Module for the MIMORE tool, which is used to investigate the morphosyntactic variation of Dutch dialects. After obtaining my Master's degree, I worked on various projects; for instance, I was a part-time lab manager at the Babylab at Utrecht Institute of Linguistics (Utrecht University), and I also worked for the AnnCor project – a project to make CHILDES, a large collection of corpora of child language, syntactically searchable. In 2016, I started my current position, which is a four-year PhD track at the University of Leuven, so I'm halfway through now.

### 2. How did you start collaborating with CLARIN DLU? Could you briefly describe the project you're currently involved in, called Quantity and Quality in Linguistics: Reverse Dialectometry?

I started using the tools and services of CLARIN DLU when I became a PhD student at the KU Leuven (Catholic University of Leuven). In the project that you mention, I investigate the formal properties of Dutch dialects/regiolects as spoken both in Flanders and the Netherlands, for which I use tools and resources provided by the Dutch and the Flemish consortia (OpenSoNaR, MIMORE).



### 3. Why is investigating non-standard language valuable for linguistic theory?

I'd like to answer this question by exemplifying a syntactic phenomenon related to infinitives in verb clusters. In Dutch dialects, the position of the infinitival marker "te", which is the equivalent of English "to", varies in the sense that in some dialects it gets doubled in a verbal cluster (for instance, "te zitten te werken", literally "to sit to work"), while in other dialects one of the markers gets dropped either in the first ("\_ zitten te werken") or in the second position ("te zitten \_ werken"). This is a linguistic fact that you wouldn't be able to observe if you studied only the official variant of Dutch, which is kept in check by the prescriptive rules, so such empirical data from the dialects actually give you a far more complex insight into the grammatical structure of Dutch infinitives. Additionally, MIMORE also shows the geographical distribution of the variation, which then allows you to investigate other possible grammatical phenomena that are also tied to the same pattern. Needless to say, without a tool like MIMORE it would be impossible to attain such insights into the linguistic structure of the Dutch language.

### 4. Given that you are an early-stage researcher, could you share your experience of how CLARIN can support researchers who are just starting their research and career? Do you have any advice for your fellow novice researchers?

What is great about CLARIN is that it allows you to explore a wealth of data that are already collected. Whenever you start working on a research project, you normally don't have any idea what's actually going on in the linguistic data. However, using a tool like MIMORE, you can quickly start working on a topic without having to do the field work yourself, which would of course be extremely time consuming. What is more, such resources have already been parsed and annotated by experts, so this is another aspect of CLARIN that I find amazing; it allows you to start applying its tools and resources fairly quickly, even if you don't have a lot of technical skills or a computational background.

My advice is to simply start using the available tools and resources, no matter whether you're a student or a more advanced researcher. The main problem, I think, is that not many people are aware of the research possibilities that CLARIN tools and resources afford. I know that there are many young researchers who study Dutch dialects, for example, and who would greatly benefit from using the CLARIN resources like the MIMORE databases, which are a goldmine of data. In this respect, these databases are very valuable since they not only consist of Dutch data as spoken in the Netherlands, but also all the Flemish dialects.

### 5. What can the Flemish consortium offer researchers working in the generative tradition?

I believe that GrETEL,<sup>6</sup> which is a tool that is developed by the Flemish consortium, has proven itself to be a very valuable service for a generative grammarian. What GrETEL primarily does is it allows you to efficiently search for specific syntactic constructions in the MIMORE databases without having to rely on technical knowledge about complex query languages. Normally, you would have to spend a lot of time searching a database for all the variants of a specific syntactic construction, but with GrETEL you only input an example of your own that conforms to the syntactic pattern you're interested in, and you immediately get all the relevant data.

### 6. Is working with a research infrastructure an established practice in your research community?

Well, it depends on what you consider my research community. As far as my fellow PhD students are concerned, a lot of them indeed make use of research infrastructures. Within the formal framework, however, researchers use it infrequently. Though I don't think it should be obligatory for generative grammarians to use corpus data in all their work, I still believe that many non-empirical researchers would still find it very helpful if they checked their claims in corpora. For one thing, corpus data can show that your intuition about a linguistic phenomenon isn't really all that representative across dialects. On the other hand, using a corpus-based approach can provide a very good stepping stone for a beginner, since it quickly shows you what the relevant linguistic situation looks like, from which you can then move on to making formal claims.

<sup>6</sup> <https://www.clarin.eu/showcase/gretel-search-engine-querying-syntactic-constructions-treebanks>



**7. Since you are also involved in teaching, which is a major priority of CLARIN's user involvement initiative, can you tell us how you integrate the resources and tools provided by the Flemish consortium in your courses? Do you have any suggestions how the link between CLARIN and university curricula could be strengthened?**

I teach a Master's course on syntactic variation in Dutch and Flemish dialects with my supervisor. We usually spend one class showing the students how to use MIMORE. We ask them to pick a specific formal topic or problem and then investigate how the formal claims correspond to the data in MIMORE. We then show the students how to use such data in their writing assignments. The main goal of the course, which I think is an important one, especially in terms of bridging the gap between the empirical and formal worlds, is how to work with a large dataset (MIMORE is comprised of 267 dialects) and apply the empirical data to a formal analysis, which is a far from trivial problem.

In the end, students really like this approach because it often allows them to get fairly novel results, even at the beginning stages. And many syntactic topics would be impossible to tackle were it not for these tools. We also use GrEtel to help students write their Bachelor's and Master's theses.

As for university curricula in general, I think the main problem is that only a few teachers have experience of combining both worlds – that is, formal analyses with empirical research. I think the first step that must be made is to encourage professors, post-docs and PhD students who also teach to become aware of these tools and resources and show them how to implement them in their courses. I think that guidelines like the Educational Module could really help in this regard.

**8. Given that you have experience with two CLARIN consortia (apart from CLARIN DLU, you previously worked with CLARIN-NL), could you describe how the two complement each other?**

What must be understood is that the division between the Dutch speaking part of Belgium and the Netherlands is a political state of affairs that does not correspond to the division of dialects. That is, dialects do not know political borders. However, tools and resources like MIMORE and GrEtel, which were often developed in collaboration by researchers working with both consortia, also contain data from Flemish dialects along with the data from Dutch as spoken in the Netherlands.

Consequently, such tools which transcend borders in this sense are really the only way to get an accurate linguistic representation of our language, and it's for this very reason that in our courses/thesis supervision at the KU Leuven we use both GrEtel, which is "our" tool, and MIMORE, which was developed by CLARIAH-NL. Additionally, the Educational Module that showcases MIMORE and GrEtel is still being updated by Sjeef Barbiers, who is from the University of Leiden, and Ineke Schuurman and Liesbeth Augustinus, who work at the KU Leuven Leuven, which I think is a great cross-border collaboration. Consequently, I see no reason why other consortia should not also collaborate in a similar manner, especially if the languages in question are similar.

**9. What would you say is the first thing CLARIN should do to be even more useful for researchers in your field?**

I would find it really wonderful if researchers could use tools like GrEtel and MIMORE to search for historical variants of Dutch (dialects). I also know that there is a lot of dialect material that cannot be accessed yet, which is something that I would like to see available through CLARIN one day, but I understand that this is often related to copyright problems.



Bruges, Belgium | photo by Waldo Miguez | Pixabay

## COLOPHON

*This brochure is part of the 'Tour de CLARIN' volume I (publication number: CLARIN-CE-2018-1341, November 2018).*

### **Coordinated by**

Darja Fišer, Jakob Lenardič and Karolina Badzmierowska

### **Written by**

Catia Cucchiarini, Griet Depoorter, Katrien Depuydt, Ineke Schuurman, Leen Sevens, Hans Westgeest, Darja Fišer and Jakob Lenardič

### **Edited by**

Darja Fišer and Jakob Lenardič

### **Proofread by**

Paul Steed

### **Designed by**

Karolina Badzmierowska

### **Online version**

[www.clarin.eu/Tour-de-CLARIN/Publication](http://www.clarin.eu/Tour-de-CLARIN/Publication)

### **Publication number**

CLARIN-CE-2018-1341  
November 2018

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International Licence.



### **Contact**

CLARIN ERIC  
c/o Utrecht University  
Drift 10, 3512 BS Utrecht  
The Netherlands  
[www.clarin.eu](http://www.clarin.eu)





