The *European Network of e-Lexicography* (ENeL)

Tanneke Schoonheim

On October 11th 2013, the kick-off meeting of the *European Network of e-Lexicography* (ENeL) project took place in Brussels. This meeting was the outcome of an idea ventilated a year and a half earlier, in March 2012 in Berlin, at the *European Workshop on Future Standards in Lexicography*. The workshop participants then confirmed the imperative to coordinate and harmonise research in the field of (electronic) lexicography across Europe, namely to share expertise relating to standards, discuss new methodologies in lexicography that fully exploit the possibilities of the digital medium, reflect on the pan-European nature of the languages of Europe and attain a wider audience.

A proposal was written by a team of researchers from the Instituut voor Nederlandse Lexicologie, the Fryske Akademy and the Berlin-Brandenburgische Akademie für Wissenschaften, partners from all over Europe were found and in May 2013 the action was approved of by COST (European Cooperation in Science and Technology), which is an EU framework supporting cooperation among scientists and researchers across Europe. This means that up to the end of 2017 public funds are available for meetings, workshops, training schools and conferences with regard to the European Network of e-Lexicography.

This presentation of ENeL explains the background, aim, structure and impact of the network, and its meetings, workshops, training schools and scientific missions.

General background

The background for setting up the network was that computers and the internet have seriously changed the conditions for the production and reception of dictionaries. The internet offers entirely new possibilities for developing and presenting dictionary information, such as with the integration of sound, maps or video, and various novel ways of interacting with dictionary users. For editors of scholarly dictionaries the new medium is not only a source of inspiration, it also generates new and serious challenges that demand cooperation and standardization on various levels:

a. Through the internet scholarly dictionaries can potentially reach large audiences. However, at present scholarly dictionaries providing reliable information are often not easy to find and are hard to decode for a non-academic audience; 'traditional' dictionary users tend to use easily accessible non-academic and user-generated dictionaries. In order to let a larger audience benefit from higher quality dictionary content, we should bridge the gap between the general public and scholarly dictionaries by improving access to these dictionaries and making them more widely known.

b. Most scholarly dictionary projects make their products available on the internet or have plans for going online. All of them find themselves confronted with similar problems relating to technologies for producing lexicographical content, presentation, interaction with users, etc. Most dictionaries take different approaches to and find different solutions for these problems. So far European and international cooperation in these fields has been restricted to bilateral collaborations. There is a clear need for a broader and more systematic exchange of expertise and for the establishment of common standards and solutions.

c. In the past years, innovative forms of electronic dictionaries

Editor | **Ilan Kernerman**

**Joint Symposia of
JACET Society of English
Lexicography and
Kansai English
Lexicography Circle
30 May 2014
Aichi University, Nagoya**

The joint symposia of the JACET Society of English Lexicography and the Kensai English Lexicography Society (KELC) was organized by Professors Michihisa Tsukamoto and Yoshihito Kamakura from Aichi University in Nagoya.

If Japan is counted among lexicographic powerhouses, these organizations are two of the driving forces behind it, contributing to the improvement and promotion of theoretical and practical lexicography.

JACET Society of English Lexicography is a group under the auspices of the Japan Association of College English Teachers. It was co-founded in 1995 by Professors Minoru Murata and Kosei Minamide, and chaired by Professor Murata from 1995 to 2006 and Professor Kaoru Akasu from 2007 to 2013. The society holds an annual workshop (including a joint one with its sister society JACET Vocabulary Acquisition Research Group once in three years), and has played a key role in organizing the two Asialex conferences held in Japan (Tokyo 2003 and Kyoto 2011). KELC was founded in 2003 with a view to read literature

▶

have appeared that no longer resemble traditional paper dictionaries but try to fully exploit the new possibilities of the digital medium. Though serious attempts have already been made at embedding electronic lexicography into a theoretical framework, an up-to-date research paradigm and common standards for electronic lexicography are still lacking, as are common standards and cooperation for interlinking content of digitized dictionaries and innovative e-dictionaries.

d. The digital medium offers the potential for a new type of lexicography that no longer views languages as isolated entities. Language migration has always been part of human history, yet this is often not properly reflected in dictionaries. Many dictionaries have their origins in the 19th century, where a national perspective on language prevailed. Consequently, the information in the dictionaries is mostly presented from the point of view of a single language, not sufficiently taking into account similar or related developments in other European languages. This view of languages often dominates also modern dictionaries. There is a clear need for a common approach to e-lexicography that forms the basis for a new type of lexicography that fully embraces the pan-European nature of much of the vocabularies of languages spoken in Europe.

Aim of the network

The aim of the European Network of e-Lexicography is to increase, co-ordinate and harmonise research in the field of e-lexicography and to make authoritative information on the languages of Europe easily accessible.

The network will:

- Make lexical knowledge of small and large languages available in a European dictionary portal. This portal will serve as the central reference point for all dictionary users who look for reliable, authoritative dictionary information on the languages of Europe and their histories on the internet.
- Enable cooperation and the exchange of resources, technologies and experience in e-lexicography and provide support for dictionaries that are not yet online.
- Discuss and aim at establishing standards for innovative e-dictionaries that fully exploit the possibilities of the digital medium.
- Establish new ways of representing the common heritage of European languages by developing shared editorial practices and by interconnecting already existing information.

Working groups

Four working groups are the main vehicles of delivering the scientific program. Each consists of lexicographers and computational linguists from different countries and includes both experienced researchers and young ones. The working groups are led by a chair and vice-chair, who also provide annual reports on its activities.

All researchers affiliated to an institution in a COST country can join the network. There is also room for researchers from countries neighbouring a COST country and from Argentina, New Zealand and South Africa.

**Working group 1.
Integrated interface for European dictionary content**

Working group 1 investigates how authoritative dictionary information on the languages of Europe can be made accessible to both the general and academic public. Chair of the working group is Anne Dykstra (Netherlands) and vice-chair is Bob Boelhouwer (Netherlands).

The working group will:

- Set up a European dictionary portal, which will give information on scholarly dictionaries of the languages of Europe and provide access to these dictionaries. Different parameters will be considered, e.g. dictionary type, language covered, digitally-born versus retro-digitized.
- Investigate the possibilities of interlinking the contents of European dictionaries.
- Investigate user requirements with respect to the presentation of dictionary content.
- Investigate the possible involvement of users in the creation of dictionary content.

**Working group 2.
Retro-digitized dictionaries**

In working group 2 the focus is on the digitization of paper dictionaries. It intends to set up guidelines and standards for turning paper dictionaries into digital format. Chair of the working group is Vera Hildenbrandt (Germany) and vice-chair is Vladimír Benko (Slovakia).

Work will be carried out on:

- The development of standards for encoding of information and the description of relevant categories for print dictionaries.
- Presenting an overview of software for the conversion of physical layout information to logical information.
- The investigation of relevant information categories to be added to the dictionary in order to make its content more readily accessible and interoperable.
- The development of a work plan for

digitisation, including parameters necessary for estimating costs.

- The investigation of possible use of dictionary content for computational linguistic applications.
- The organisation of a training school on standard tools and methods for retro-digitising dictionaries in 2015.

Working group 3.
Innovative e-dictionaries

The scope of working group 3 is the development of digitally-born dictionaries, focusing in particular on the latest developments in e-lexicography and the interface between lexicography and computational linguistics. Chair of the working group is Simon Krek (Slovenia) and vice-chair is Carole Tiberius (Netherlands). Work will be carried out on:

- The description of the workflow for corpus-based lexicography.
- Providing an overview of existing software needed to set up this workflow.
- The use of dictionary writing systems.
- The analysis of the possible impact of automatic acquisition of lexical data.
- The analysis of the interface between dictionary and computational lexica (cf. wordnets) and syntactically and semantically annotated corpora (cf. FrameNet, SemCor, Senseval).
- The investigation of the possible use of dictionary content for language technology applications.
- The organisation of a training school on innovative approaches in e-lexicography in 2016.

Working group 4.
Lexicography and lexicology from a pan-European perspective

Working group 4 investigates how the pan-European nature of the vocabularies of the languages of Europe can be represented in single-language dictionaries and within the European dictionary portal. This is particularly relevant to studying the multiple dimensions of borrowing, i.e. the migration and re-migration of words and meanings across the languages of Europe. Chair of the working group is Eveline Wandl-Vogt (Austria) and vice-chair is Phil Withington (UK).

This working group will:

- Develop ways in which already existing information from single language dictionaries can be displayed and interlinked to represent more adequately their common European heritage.
- Develop editorial guidelines for the integration of European information into more traditional dictionaries as well as innovative ones.

- Find new applications for the very large amount of interconnected dictionary information from the European dictionary portal in the field of digital humanities.
- The organisation of a training school in lexicography and lexicology from a pan-European perspective in 2017.

Training schools and short term scientific missions

One of the responsibilities of the network is to organise training schools and to provide so-called short-time scientific missions (STSMs). The three training schools mentioned above will be organised by members of the relevant working groups, in close cooperation with the training school manager of the network, Rute Costa (Portugal). Participants of the training schools will get their travel and accommodation expenses reimbursed by the network.

The first training school will be held in 2015 on the subject of retro-digitization. More information on this training school will be available at the time of the network meeting in July 2014 and will also be published on the website: http://www.elexicography.eu/events/training-schools/.

As part of an STSM researchers visit established dictionary projects and centres of excellence in a country other than their own. The visit period may take from five days to three months and can range from discussions and demonstrations to deeper involvement in the activities of the centres visited, in accordance with the needs of the researchers. Travel and accommodation expenses of approved missions are reimbursed by the network. There is room for at least four STSMs in each year, which brings the total of STSMs to take place within the duration of the project (till October 2017) to sixteen or more.

STSMs are especially meant to allow young researchers to build up their own networks, and they will facilitate and increase the capacity for research in the field of e-lexicography. Tanneke Schoonheim (Netherlands) is appointed as the manager of these missions. For more information on the missions and how to apply, see http://www.elexicography.eu/events/workshops/.

Impact of the network

The network will allow the exchange of knowledge and expertise in the field of e-lexicography. It is open to all scholarly dictionaries in Europe irrespective of their previous experience with e-lexicography. The network will facilitate the coordination and progressive expansion and standardization of research activity through the work of the four working groups and

▶

about lexicography, lexical semantics and corpus linguistics, and deepen the understanding of recent research findings in these fields of study. Coordinated by Professor Kosei Minamide, the circle meets bi-monthly to discuss the issues and insights raised in selected books, and its members have translated and brought three of these to publication in Japan: *Phraseology* (Cowie, ed. 2001), *Words and Phrases* (Michael Stubbs, 2001), and *Lexicography: An Introduction* (Howard Jackson, 2002). The regular meeting of the JACET group is in Tokyo and that of KELC is in Kyoto. Located between the two cities, Nagoya was a fine venue for the joint symposia. The meeting began with a special session on 'Challenges Facing English Dictionary Editors', with the participation of the chief editors of four major English-Japanese learners' dictionaries, and continued with sessions on 'EFL Dictionaries: Description and Use' and 'Collocation and English-Japanese Dictionaries'. Altogether there were a dozen talks, attended by 80 participants. The symposium included a book display and was followed by a buffet-style party at the university cafeteria.

Shigeru Yamada
Chair, JACET Society of English Lexicography

through close cooperation with the Euralex and e-Lex conferences.

The sustainability of the network will be promoted through the European dictionary portal and through the publication of reports and roadmaps concerning future standards for e-lexicography on the network's website. This website is intended to become the first source of consultation by lexicographers planning to create new e-dictionaries.

In the European dictionary portal reliable information on the languages of Europe will be made easily available to everyone. The portal will therefore be a trustworthy alternative to the many user-generated dictionaries on the internet. In addition, whereas traditional printed dictionaries are often inaccessible to people with a visual impairment (set in very small type, heavily abbreviated for space-saving reasons) the portal will give visually impaired people easy access to these dictionaries with screen readers.

The development of common standards will save time and money. New e-dictionaries or dictionaries that are due to go online will no longer have to develop their own standards but will be able to refer to the publications of the network. In particular dictionaries of small languages will benefit from such common and easily-available standards, as they usually do not have the means to develop their own approaches. The network can therefore substantially contribute to regional or minority languages.

The overall quality of the dictionaries will be improved. The new editorial methods and practises to be developed will reflect more realistically the language situation in Europe and the historical development and interaction of European languages, in particular the migration of words and meanings across Europe. Lexicographers and lexicological researchers, especially young ones, will increase their knowledge and skills by participating in working groups, training schools and conferences. Yvonne Luther (Germany) is appointed as manager for young researchers and/or female researchers to make sure that they get the right chances in the network.

The large amount of data connected in the European dictionary portal will enable new lines of research in the field of digital humanities that could have not been carried out on the basis of isolated language resources, e.g. the spread of technological innovation by studying the appearance of relevant words in the vocabularies of the languages in Europe.

Structure of the network

The network is driven by a management committee, comprising an elected chair and vice-chair and up to two representatives from each COST country. The chair is Martin Everaert (Netherlands) and vice-chair is Iztok Kosem (Slovenia). The management committee meets every six months and is responsible for co-ordinating the activities of the working groups, budget planning and the allocation of funds, organising training schools and conferences. It will monitor progress in relation to the scientific focus and work plan in relation to the achievement of milestones.

The steering group is responsible for preparing annual reports on the work of the network, overseeing the development and maintenance of the network website, communicating with the COST office and monitoring its procedures. It comprises the chair and vice-chair of the management committee, the chair and vice-chair of each working group, and the managers for the training schools, STSMs and young/female researchers. The steering group meets every three months, having started in January 2014 in Leiden, followed in April in Vienna, in July in Bolzano and in October in a place yet unknown.

At present the network comprises representatives from 29 COST countries: Austria, Belgium, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Latvia, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland and the United Kingdom.

Members of the network meet every six months, preferably in combination with other conferences regarding lexicography or computational linguistics, such as Euralex, e-Lex and LREC. On these occasions there will always be meetings of the working groups, the management committee and the steering group, followed by a plenary session for all members on the progress made. The next network meeting is on July 19-20 in Bolzano (Italy), directly following the Euralex conference. The first meeting of 2015 is due in Vienna (Austria) in January or February and the second meeting may be connected to e-Lex 2015 in July in Herstmonceux Castle (UK).

For more information on ENeL, please visit our website: www.elexicography.eu.

EUROPEAN NETWORK OF e-LEXICOGRAPHY