**Subject | Minutes of the 2nd Meeting of the Steering Group of COST Action IS1305 “European Network of e-Lexicography (ENeL)”**

**at Vienna, Austrian Academy of Sciences,**

**Dr. Ignaz Seipel-Platz 2, 1010 Vienna (Austria)**

**on 14 April 2014, from 9h00 till 17h00**

Present: Martin Everaert (Chair), Iztok Kosem (Vice-Chair), Anne Dykstra (WG1 Chair), Bob Boelhouwer (WG1 Vice-Chair), Vera Hildenbrandt (WG2 Chair), Vladimir Benko (WG2 Vice-Chair), Simon Krek (WG3 Chair), Eveline Wandl-Vogt (WG4 Chair), Rute Costa (Training Schools Manager), Tanneke Schoonheim (STSM Manager)

Absent with notice: Carole Tiberius (WG3 Vice-Chair), Phil Withington (WG4 Vice-Chair), Yvonne Luther (ESR/Female Researcher Manager)

1. **Opening and Welcome**

Martin Everaert (Chair of the Action) opened the meeting and welcomed the participants.

1. **Action list Leiden and Announcements**

* All topics on the Leiden action list were on the agenda of this meeting and were discussed
* On 5 and 6 May all Chairs of the running COST actions will meet in Malta. Martin will attend this meeting where presentations of various actions will be given. The Progress Management Report that is requested for this meeting will be sent to the members of the Steering Group.
* In the Progress Management Report two projects/applications are mentioned which might be funded with additional money thanks to our COST action. One is a Slovene-Belgian project on neologisms and the other is a Swiss project about interlinking dictionary content to other resources. There will be short presentations of these projects at the MC meeting in Bolzano. It is important to point out these possibilities to the participants of the action.

1. **Working Groups**

The action has four main objectives, for each Working Group (WG) there is a special focus on the subject of European e-Lexicography. There is a common understanding that all WGs have to cooperate closely and that the requirements of the participants have to be coordinated to avoid gaps or double work. It is important that all participants of all WGs contribute as much as possible. The Chairs of the WGs will make plans and deliverables for the meeting in Bolzano.

The portal we want to develop will be the common result of the action, containing the results of all WGs. It seems wise to make a difference between results that are meant for the general public and results that are specifically meant for a public of experts (e.g. lexicographers, computational linguists, historians, philosophers, lawyers, translators and other people with an interest in language and linguistics). This implies that there might even be two portals, one for the general public and one for the expert public. WG1 will set realistic standards for these portals and investigate the amount of linking that can be done. WG1 will also investigate user requirements and user involvement in the portals. The other WGs will cooperate in this, specifying in what way they want to show their results in each of these portals.

Annexes 1 (WG1) and 2 (Features for the Dictionary Portal) are discussed and the elements that can be realised within the action will be used for further discussion of WG1 in Bolzano. Because of the complexity of the subject, there will be no further discussion on topics such as search in single dictionaries, federated search in multiple dictionaries, search in source language while getting results in target language, interlinking on the level of meanings, and the possibilities for use of the portal by machine interfaces, but these can be put on a list of possible follow-up projects. The same goes for the possibility to add national corpora to the portal.

The expert portal will contain:

* information on the national dictionaries of the status languages of the European countries involved in the action, including languages that might not be in active use any more (e.g. Latin, Gothic). For the inventory of these dictionaries the MC will be consulted. Dialect dictionaries and etymological dictionaries are not to be excluded a priori.
* access to a limited amount of these dictionaries (i.e. the scholarly or near-scholarly dictionaries of these languages that are freely available online only). WG1 will decide which dictionaries to include, and what amount of interlinking is realistic. WG2 can base the work flow of the topic of retro-digitisation on the requirements of WG1 for the dictionaries to be included in the portal.
* information on and guidelines for standards in the field of e-lexicography.
* description of ongoing projects and events in the field of e-lexicography.
* publication of the results from case studies in the field of e-lexicography that have been done by all four WGs.
* blogs and discussion lists.

The general portal will contain:

* information on the national dictionaries of the status languages of the European countries that are in the general portal (i.e. scholarly or near-scholarly dictionaries of these languages that are freely available online only).
* access to these dictionaries. WG1 will decide in what way and to what amount they can be accessed.
* publication of the results from case studies on the topic of e-lexicography that have been done by all four WGs and that are of interest to the general public.
* blogs and discussion lists.

We also have to think about strategies to make both portals more prominent for the intended public. A page on Facebook will be made by Anne, and a Twitter account will be activated by Simon. Furthermore, in order to attract the expert public, workshops and meetings are connected to big events such as EURALEX, eLex, LREC and so on.

One of the objectives of all WGs is to produce a road map to possible ways of linking and interconnecting data in the future and to generate new lines of research in the field of digital humanities. It is important to give language a place in the field of cultural heritage.

Annex 3 (WG2) is discussed. WG2 will make an inventory of dictionaries that have been and are to be retro-digitised with information on dictionary type, source and target language, format and structure etc., and publish this inventory on the website. An overview of software for the conversion of physical lay-out information to logical information is also one of the contemplated results. The MC and WG members will be asked to provide this information. In order to reach this objective, it is important to acquire DTDs or XML-schemas of the dictionaries involved.

The process of retro-digitisation will be described with the focus on the main problems that can occur during it. It is important to write guidelines in such a way that policy makers understand them and are willing to spend money on such projects. *Development of a standard workflow for digitisation* and *Standards for the encoding of information and the description of relevant information categories for print dictionaries* are the first objectives to be tackled in the first and the second year of the action.

The development of concepts for linking to other resources (for instance to WordNet or other types of lexical information) will be tackled in cooperation with WG1 and WG3. It is important to know what has already been done and what is currently planned in the participating countries.

The search for future funding for retro-digitisation is first and foremost the responsibility of the Steering Group. It is important to make governments aware that they have to invest in lexicography because most commercial companies are no longer able to do so.

Annex 4 (WG3) is discussed. Innovative dictionaries are dictionaries which are 1) digitally born, or 2) retro-digitised and enlarged with digital features. An innovative dictionary is not supposed to be printed as such, although parts of it may be printed for special purposes. It is not so much the content of the dictionary that has to be innovative, but more the way in which the content is presented.

The topics of WG3 are described in the annex. In year 1 of the action the work on topic 1 (description of the workflow for corpus-based lexicography), topic 2 (overview of existing software needed in this workflow) and topic 3 (Dictionary Writing Systems (and Corpus Query Systems)) will start. The description of the workflow is on the agenda for the WG meeting in Bolzano. The results of the research on the topics will result in deliverables, to be presented at various conferences (EURALEX, eLex etc.). Work on the other topics doesn’t have to wait until the topic is on the agenda. It is important to make room at the meetings for the presentations on these topics as well. Topic 6 (Investigation of possible use of dictionary content for computational linguistic applications) will be dealt with in cooperation with WG2 and WG4. The topics are targeted at lexicographers and computational linguists, not for the broader public. This WG is meant especially for uniting the specialists. There are already volunteers in WG3 who will be responsible for the organisation of a meeting on the topic of their choice.

Annex 5 (WG4) is discussed. Languages can be seen as cultural heritage. It may be interesting to investigate the languages of Europe in the context of Eurolinguistics: how do languages behave in mutual contact? It is good to provide broader access to (etymological) information from dictionaries in addition to what is already available on the internet (Wikipedia, Wiktionary etc.). Tasks of WG4 are to develop standards on interlinking words on a semantic level and to set standards on encoding etymological information to be able to link this information to other dictionaries. Questions that have to be answered are how to give unique identifiers to words from different languages who share the same origin. The standard terminology used in the deliverables of WG4 (case studies on various topics, such as sports, food and drink, household furniture etc.) also needs to be set.

It might be interesting to investigate how we can make the information that is hidden in dictionaries usable for projects in other disciplines.

1. **Dissemination**

The subject of dissemination is on the agenda for Bolzano. It is important to reach both the specialist and the general public. For the specialist public there will be deliverables on the website (inventories, bibliographies etc.), and notifications of congresses and other meetings and publications in journals. For the general public it is important to search for interesting topics; before the World Cup in Brazil will start (June 2014), we will try to produce an article and a website about the words used in the game of football in Europe, to show to what extend the original English words have been adapted into other European languages and if and how these languages have come up with their own terminology in this field.

The results of WG3 on innovative dictionaries (due for 2017) can be used in the handbook on this subject that Routledge has offered to publish.

1. **Training Schools**

Annex 6 (Training Schools) is discussed. The Training Schools are meant for Early Stage Researchers and advanced students, who will get a reimbursement for attending the school. Experienced MC and WG members will teach at these training schools and will be reimbursed. The budget for the Summer School is not allocated yet, but will be at least 25.000 euro. We will aim to connect Training Schools with existing Summer Schools. To every Training School a workshop for WG members can be added, maybe even a MC meeting, in order to save costs.

All MC members will be asked for information on summer schools that can host one of the Training Schools of the action. The Training School of 2015 will be on the topic of retro‑digitising. It can be held in Lisbon (the final decision will be made at the MC meeting in Bolzano). It will last one week of between 15 and 45 teaching hours and can contain as many as 20 students. Vera and Vlado will make a program and look for teachers.

The locations of the Training Schools of 2016 (Innovative e-dictionaries) and 2017 (Pan-European Heritage) will be discussed by the MC in July in Bolzano.

1. **Short Term Scientific Missions**

The call for Short Term Scientific Missions (STSMs) will be put on the website soon. It will also be sent to relevant mailing lists. STSMs are open for all researchers. For this year there are at least three STSMs available that have to be completed before October 10th. The SG will decide on assigning the requested STSMs. There are two dates for deciding on STSMs, June 15th and August 15th.

1. **Website**

The website of the action is made by Iztok Kosem ([www.elexicography.eu](http://www.elexicography.eu)). All WGs are asked to provide as much information as possible. It is an open website, for which no password is needed. This means that it will not contain unfinished documents. These are to be spread via the mailing lists in Google Groups.

1. **Communication**

Google Groups will be used for all communication. The Chairs of the WGs will make a group for their WG and the members of the SG will be on all of them, to be informed about every aspect of the action.

1. **AOB**

There is no AOB.

1. **Closing**

Martin closes the meeting and thanks the participants for their presence and cooperation.