



**European Cooperation
in the field of Scientific
and Technical Research
- COST -**

Brussels, 24 May 2013

COST 029/13

MEMORANDUM OF UNDERSTANDING

Subject : Memorandum of Understanding for the implementation of a European Concerted Research Action designated as COST Action IS1305: European Network of e-Lexicography (ENeL)

Delegations will find attached the Memorandum of Understanding for COST Action IS1305 as approved by the COST Committee of Senior Officials (CSO) at its 187th meeting on 15-16 May 2013.

MEMORANDUM OF UNDERSTANDING
For the implementation of a European Concerted Research Action designated as
COST Action IS1305
EUROPEAN NETWORK OF E-LEXICOGRAPHY (ENeL)

The Parties to this Memorandum of Understanding, declaring their common intention to participate in the concerted Action referred to above and described in the technical Annex to the Memorandum, have reached the following understanding:

1. The Action will be carried out in accordance with the provisions of document COST 4154/11 “Rules and Procedures for Implementing COST Actions”, or in any new document amending or replacing it, the contents of which the Parties are fully aware of.
2. The main objectives of the Action are to (a) make lexical knowledge of the languages of Europe available in a European dictionary portal, (b) enable cooperation and the exchange of resources, technologies and experience, (c) discuss standards for innovative e-dictionaries, (d) establish new ways of representing the common heritage of the languages of Europe.
3. The economic dimension of the activities carried out under the Action has been estimated, on the basis of information available during the planning of the Action, at EUR 64 million in 2013 prices.
4. The Memorandum of Understanding will take effect on being accepted by at least five Parties.
5. The Memorandum of Understanding will remain in force for a period of 4 years, calculated from the date of the first meeting of the Management Committee, unless the duration of the Action is modified according to the provisions of Chapter IV of the document referred to in Point 1 above.

A. ABSTRACT AND KEYWORDS

Computers and the availability of the World Wide Web have significantly changed the conditions for the production and reception of dictionaries. For editors of scholarly dictionaries the Internet is not only a source of inspiration, it also generates new and serious challenges: (a) To give users easier access to scholarly dictionaries and to bridge the gap between the general public and scholarly dictionaries, (b) To establish both a broader and more systematic exchange of expertise and of common standards and solutions, (c) To develop 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 the languages spoken in Europe. The proposed Action aims to establish a European network of lexicographers in order to deal with the above issues in a structured way.

Keywords: e-lexicography, e-dictionaries, corpus-based lexicography, European languages, European standards in lexicography

B. BACKGROUND

B.1 General background

Computers and the availability of the World Wide Web have significantly changed the conditions for the production and reception of dictionaries. The Internet offers entirely new possibilities for the presentation of dictionary information, for example, 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 a very large audience.

However, at present scholarly dictionaries providing reliable information are often not easy to find and hard to decode for a non-academic audience; ‘traditional’ dictionary users tend to use easily accessible non-academic, user-generated dictionaries. In order to let a larger audience benefit from high 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 are planning to go online. All of them find themselves confronted with similar problems relating to technologies for producing lexicographical content, presentation of content, interaction with dictionary users etc. Most dictionaries take different approaches to and find different solutions for these problems. So far

European 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 few years, innovative electronic dictionaries have been created 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, a new research paradigm and common standards for electronic lexicography are still lacking. And so are common standards and cooperation for the interlinking of the content of digitized dictionaries and innovative e-dictionaries.

(d) The digital medium has the potential for a new type of lexicography that no longer views languages as isolated entities. Language migration has always been part of language 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 also dominates 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 the languages spoken in Europe.

The COST framework guarantees the European nature of the proposed Action and provides the right forum for networking and capacity-building as envisaged in the Action. COST is currently the only research framework on a European level that funds large-scale humanities networking projects. Professional associations of lexicographers, which hold biennial conferences, have expressed interest in cooperation with the Action. The Action will further seek cooperation with CLARIN-ERIC (Common Language Resources and Technology Infrastructure - European Research Infrastructure Consortium), which is committed to establish an integrated and interoperable research infrastructure of language resources and their technologies and with DARIAH (Digital Research Infrastructure for the Arts and Humanities), which aims to enhance and support digitally-enabled research across the arts and humanities.

B.2 Current state of knowledge

Most scholarly dictionaries are already available on the Internet in a retro-digitized version or are planning to go online. More recently, innovative e-dictionaries are trying to exploit the new possibilities of the digital medium (de Schryver 2003), for example the *Algemeen Nederlands*

Woordenboek (Schoonheim/Tempelaars 2010) or *elexiko* (Klosa et al. 2006). At the moment, “[l]exicography is in a state of transition ... between the technology of the printed word and the bound book ... and the technology of the Internet and the electronic product” and “[i]t is too early to say what form innovative dictionaries of the future will take” (Hanks 2012: 82).

E-Lexicography

Research and future directions in e-lexicography are hot topics, and understandably so. We are undeniably in the era of online dictionaries and reference tools. Both demand knowledge of electronic lexicography and information science, which requires a reassessment of the field. Modern dictionary users expect the dictionary to apply all the new technologies that they have become used to. Online scholarly dictionaries should therefore be sophisticated and user-friendly search engines as well. Though serious attempts have already been made at embedding electronic lexicography into a theoretical framework, there still is an urgent need of a new research paradigm for electronic lexicographic mediums. The organizers of the third eLex conference (2013) would seem to acknowledge that e-lexicography is actually still in its knee breeches. Current discussions on the differences between paper and electronic dictionaries, they say, “merely take away the focus from the actual task at hand – to find different ways on how to exploit the rich potential of electronic medium in order to respond (quickly) to the needs of the new types of users, as well as to the needs of modern lexicographers, to forget about conventional approaches and be innovative, to conceptualize the dictionary with an electronic format in mind”. Their adagium is: “think outside the paper” (<http://eki.ee/elex2013/>).

User-generated dictionary content

The two-way nature of Web 2.0 has also changed the role of Internet users: they are no longer passive consumers of online content that was created by specialists, but they can also take an active role in creating content. Not surprisingly, the last few years have seen an increase of user-generated dictionaries and dictionary-content (Carr 1997, Mann 2010), the best-known of which is *Wiktionary* with its currently 145 language editions, of which 54 are languages spoken in Europe (Meyer/Gurevych 2012). In terms of numbers of users and numbers of entries these collaboratively created “bottom-up” dictionaries now often outperform online academic dictionaries. In terms of dictionary content, however, they are often heavily out-dated and of poor quality (Lew 2012). Many of them are based on out-of-copyright editions of older print dictionaries and they usually do not take into account recent research in lexicography, e-lexicography, lexicology or linguistics (Hanks 2012), like for example corpus linguistics, which is now a standard approach. We thus observe a

clear discrepancy between the concepts of theoretical lexicography and the present standards of academic (e-)lexicography on the one hand and the practice of these user-generated online dictionaries on the other hand (Mann 2010). The obvious success of user-generated dictionaries warrants careful study. Lorentzen/Theilgaard (2012) show that much can be learned from studying the behavior of users of online dictionaries. Their line of research will no doubt prove to be worthwhile for the Action (see also Müller-Spitzer et al. 2011).

The common linguistic heritage of the languages of Europe – consequences for lexicography

Historically, Europe can be regarded as a ‘common space of knowledge’ with a common background for lexicography and with shared lexicographical practices and methodological principles (Haß 2010, 2011). The big national dictionaries of the 19th century, like the *Oxford English Dictionary*, the *Woordenboek der Nederlandsche Taal*, the *Svenska Akademiens Ordbok*, the *Trésor de la langue française* and the *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm* were influenced by a common canon of dictionaries and other reference works and must be regarded as adaptations of the already existing dictionaries rather than as independent works (Haß 2010, 2011; Kirkness 2012). At the time of their creation one important purpose of these dictionaries, however, was to contribute to nation-building, which explains the fact that the historical development and synchronic use of the languages described in such a dictionary are studied from the point of view of a single language only. This does not adequately reflect the common pan-European *linguistic* heritage (Munske/Kirkness 1996; Habermann 1999) of the languages of Europe, which in fact have been in permanent contact with each other. Therefore lexical equivalences between them are to be understood as the results of potentially ongoing long-term exchange processes, rather than the isolated events of borrowing, as which they are still viewed in most dictionaries. Simpson (2004) shows that taking a European perspective gives a more realistic picture of the actual development and is the only adequate one (Haß 2012). Görlach (2001), which records the use and spread of Anglicisms across 16 European languages, is an example of a dictionary that takes a European perspective on language change.

B.3 Reasons for the Action

Although there are examples of research in the field of e-lexicography in Europe, there is a lack of common approaches and of common standards. A European Network of e-Lexicography would enable the academic lexicographical community to co-ordinate and harmonise research in the field of e-lexicography. The COST Action aims at establishing common standards with respect to

technologies for both retro-digitized and new electronic dictionaries and at developing a joint approach to e-lexicography and new methods in lexicography that fully explore the possibilities of the digital medium and reflect the pan-European nature of much of the vocabularies of the languages of Europe. Common standards and joint approaches are beneficial to all dictionary projects, particularly to those documenting small languages, which often do not have the means to develop their own standards.

By creating a European dictionary portal with information on and access to the academic dictionaries of Europe the Action aims to make authoritative dictionary information visible for a much wider audience and easily accessible to other lexicographers, researchers from other disciplines, writers and translators, to visually impaired people and to the general public. For small language dictionaries such a portal is an opportunity to present themselves to a wide public without worrying about printing costs – thus the portal is also a means to preserve European language diversity.

In addition, Early Stage Researchers (ESRs) in the field of lexicography will benefit considerably from the establishment of a COST Action, which will offer scientific exchanges and the coordination for joint research with more experienced lexicographers. The Action will also be involved in the training of lexicographers and will be able to involve researchers at the very first stages of their careers.

B.4 Complementarity with other research programmes

The Action will complement and cooperate with, but not duplicate other European research projects.

The Action will enhance the activities of the European Research Infrastructures CLARIN (now an ERIC) and DARIAH. The COST Action will coordinate the development of dictionary-specific standards and make these compliant with the standards developed in CLARIN and DARIAH. The Action will complement the existing activities in these European infrastructure projects, and will invite relevant stakeholders such as CLARIN and DARIAH experts to the appropriate Working Group (WG) meetings.

C. OBJECTIVES AND BENEFITS

C.1 Aim

The Aim of this Action is to increase, co-ordinate and harmonise European research in the field of

e-lexicography and to make authoritative dictionary information on the languages of Europe easily accessible. The Action will (a) 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. The Action will (b) enable cooperation and the exchange of resources, technologies and experience in e-lexicography and provide support for dictionaries which are not yet online. It will (c) discuss and aim at establishing standards for innovative e-dictionaries that fully exploit the possibilities of the digital medium. In taking a more common approach to e-lexicography the Action will (d) establish new ways of representing the common heritage of European languages by developing shared editorial practices and by interconnecting already existing information.

C.2 Objectives

The objectives of the Action are as follows:

1. To coordinate research in the field of (electronic) lexicography across Europe through the establishment and progressive growth of a European Network of e-Lexicography. The growth of the network will be measured by the numbers of researchers and ESRs joining the Action, attending Training Schools and conferences.
2. To make authoritative dictionary information on the languages of Europe accessible to both the general public and to lexicographers in a European dictionary portal. The portal will list and give information on scholarly dictionaries of the languages of Europe (it will allow sorting by various parameters, e.g. dictionary type, language covered, electronic vs. retro-digitized) and provide access to these dictionaries.
3. To further the exchange of technologies and expertise in the field of online lexicography and e-lexicography in order to find common approaches and provide support to other dictionary makers. The Action will produce (a) a number of reports describing standards for encoding and presentation of dictionary content, standards for user interfaces and user interaction both for the academic and non-academic user, but also covering information retrieval techniques going beyond literal search or web service development to implement interoperability of resources, (b) a roadmap towards a shared effort in developing corpus-based e-lexicography, which should pinpoint the common issues in the path from raw material to dictionary article, and result in a pooling of resources to develop the necessary technical infrastructure to support corpus-based lexicography, (c) a roadmap to possible ways for the extensive linking and interconnection of the data contained in European dictionaries in the European dictionary portal that will generate new lines of research in the field of

digital humanities.

4. To provide a foundation for the further exploration of a pan-European approach to lexicography by discussing new standards and methodologies to describe the common European heritage of much of the vocabularies of the languages of Europe and developing shared editorial practices among the dictionaries of Europe. These are to be published in a series of scholarly articles.

5. Offering support to ESRs in the field of lexicography through Short-Term Scientific Missions (STSMs), Training Schools and involvement in the Action's activities.

6. Promoting the sustainability of the network beyond the lifetime of the Action (a) through the European dictionary portal, (b) by documenting the results obtained in the Action on the Action website, which will serve as a point of reference for new electronic dictionaries or paper dictionaries that are planning to go online, (c) by identifying potential funding sources and developing collaborative funding applications.

Progress towards each Objective will be assessed in concrete terms at each Management Committee meeting during the life of the Action.

C.3 How networking within the Action will yield the objectives?

The COST Action will allow the exchange of knowledge and expertise in the field of e-lexicography; the Action is open to all scholarly dictionaries in Europe irrespective of their previous experience with e-lexicography (Objective 1). The network will facilitate the coordination and progressive expansion and standardization of research activity through the work of four Working Groups and through cooperation with relevant conferences.

Objective 2 will be achieved by WG1, in which an online dictionary portal will be designed and implemented, which will be progressively expanded during the life of the Action. WG1 will cooperate with WG2 and WG3, which are concerned with developing common standards in the field of e-lexicography (Objective 3), i.e. for retro-digitized paper dictionaries that are already online, or planning to go online (WG2) and for innovative e-dictionaries that want to exploit the new possibilities of the digital medium. These two Working Groups will work closely together in finding common approaches to linking contents of retro-digitized and innovative e-dictionaries. The results of WG2 and WG3 will be disseminated in Training Schools and will also feed back into the dictionary portal.

Objective 4 will be achieved in WG4, which is concerned with the question of how to implement a pan-European view on the vocabularies of the languages of Europe in innovative e-dictionaries. WG4 will further consider how the extensive interlinking of information from many different

dictionaries (implemented by WG2 and WG3 together with WG1) allows lexicographers to get away from a single language approach to lexicography and will investigate how these vast amounts of language data can generate new lines of research in the field of digital humanities.

Each Working Group will comprise lexicographers from different COST countries and will include both more experienced researchers and ESRs. The network will offer further support to lexicographers at very early stages of their careers by involving them in all activities of the Action (Objective 5). This will give them the opportunity to establish their own partnerships. ESRs will particularly benefit from participating in Short-Term Scientific Missions, from participating in WG activities and from attending Training Schools and Action conferences.

The sustainability of the network beyond the life of the Action (Objective 6) will be promoted through the European dictionary portal, through the publication of reports and roadmaps concerning future standards for e-lexicography on the Action website, which is intended as the first source to be consulted by lexicographers intending to create new e-dictionaries, and which will be available beyond the life of the Action. Participating in the COST Action will further enable lexicographers to extend their research networks and initiate new funding applications.

C.4 Potential impact of the Action

The Action will be beneficial on various levels:

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.

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 European dictionary portal will give visually impaired people easy access to these dictionaries (with screen readers).

The development of common standards in e-lexicography will save time and money: New e-dictionaries or dictionaries that are planning to go online will no longer have to develop their own standards but can 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 Action can therefore substantially contribute to saving regional or minority languages by including dictionaries of languages protected under the European Charter for Regional or Minority Languages.

The Action aims to improve the overall quality of dictionaries: The new editorial methods and practices to be developed during the life of the Action 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, particularly ESRs, will increase their knowledge and skills by participating in Training Schools and Action conferences.

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

C.5 Target groups/end users

The new standards and methodologies that will be developed by the Action are beneficial to the European lexicographic community and to the scientific community worldwide. In particular dictionary projects documenting small languages will benefit from accepted common standards and methods in e-lexicography. The Action is beneficial to ESRs in the field of lexicography, who will be involved in all activities of the Action and in the management of the Action. They will be able to make contact with leading experts in the field and establish their own research networks. The benefits of the new European dictionary portal with information on and access to the scholarly dictionaries of Europe extend far beyond the scientific community: the new portal will be used by linguists and researchers from other disciplines, but also by language students and students from neighbouring disciplines. Furthermore it will give people with visual impairment access to dictionaries which they could formerly not use.

The portal will also be used by people working with languages, e.g. translators, proof readers, journalists or writers and more generally, by anyone who is looking for authoritative lexical information on the languages of Europe.

D. SCIENTIFIC PROGRAMME

D.1 Scientific focus

The Action will harmonise and systematise existing and future research by developing common pan-European research standards and solutions in the field e-lexicography. Apart from the implementation of a European dictionary portal with information on and access to all dictionaries the research coordination in the ENEL Action focuses on the following areas:

(a) setting up guidelines and standards for turning paper dictionaries into a digital format,

(b) mapping current and possible future trends for the creation of born-digital dictionaries, focusing on the latest developments in e-lexicography and the interface between lexicography and computational linguistics,

(c) obtaining and overview of how the accessibility of authoritative dictionary information on the languages of Europe can be improved for both the general and academic public,

(d) exploring the possibilities of extensive linking of dictionary content from different European languages,

(e) developing shared editorial standards and discussing new methodologies to describe the common European heritage of much of the vocabularies of the languages of Europe.

Through Action conferences, Working Groups, regular plenary sessions and Short-Term Scientific Missions (STSMs) the Action will facilitate the development of multi-professional research collaborations.

The Action conferences and some of the WG meetings will take place in connection with conferences on (electronic) lexicography, when possible.

D.2 Scientific work plan methods and means

WORKING GROUPS

Four Working Groups will be the main vehicles of delivering the scientific work plan. They will focus on the coordination of different aspects of e-lexicography:

WG1 Integrated interface to European dictionary content: This WG will focus on how authoritative dictionary information on the languages of Europe can be made accessible to both the general and academic public. It will:

- Set up a European dictionary portal, which will give information on scholarly dictionaries of the languages of Europe (different parameters will be considered, e.g. dictionary type, language covered, electronic vs. retro-digitized) and provide access to these dictionaries through the portal website
- Look into the possibilities of interlinking the contents of European dictionaries
- Map user requirements with respect to the presentation of dictionary content
- Discuss the possible involvement of users in the creation of dictionary content
- Identify future funding sources and develop collaborative funding applications

There will be close collaboration with the other three WGs with respect to user-interfaces to the dictionary data.

WG2 Retro-digitized dictionaries: In this WG the focus is on the digitisation of paper dictionaries. This WG intends to set up guidelines and standards for turning paper dictionaries into digital format. Coordination work will be carried out on:

- Standards for the encoding of information and the description of relevant information categories for print dictionaries
- Software for the conversion of physical lay-out information to logical information
- Mapping of relevant information categories to be added to the dictionary in order to make the dictionary content more readily accessible and interoperable
- Development of a work plan for digitisation, including parameters necessary for estimating costs
- Discussing the possible use of dictionary content for computational linguistic applications
- Identifying future funding sources and developing collaborative funding applications

In this WG, requirements from WG1 concerning the linking of information between dictionaries and concerning user interfaces will be taken into account. In cooperation with WG4 the extent to which new insights in the lexicographical description of European language vocabulary can be taken into account in retro-digitisation projects will be discussed.

WG3 Innovative e-dictionaries: This WG will coordinate the development of born-digital dictionaries, focusing on the latest developments in e-lexicography and the interface between lexicography and computational linguistics. Work will be carried out on:

- Description of the workflow for corpus-based lexicography
- Overview of existing software needed to set up this workflow
- Dictionary writing systems
- Analysis of the possible impact of automatic acquisition of lexical data (distributional thesauri etc.)
- Analysis of the interface between dictionary and computational lexica (cf. wordnets) and syntactically and semantically annotated corpora
- Investigation of possible use of dictionary content for computational linguistic applications
- Identifying future funding sources and developing collaborative funding applications

In this WG, requirements from WG1 dealing with linking information between dictionaries and with the user interface will be taken into account. Interaction will also take place with WG4 to be able to take into account the new insights into the lexicographical description of the vocabularies of the different European languages.

WG4 Lexicography and lexicology from a pan-European perspective: This WG will discuss 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. The WG will look into:

- Developing ways in which already existing information from single language dictionaries can be displayed and interlinked to represent more adequately their common European heritage
- Developing editorial guidelines for the integration of European information into more traditional and into innovative e-dictionaries
- Finding new applications for the very large amount of interconnected dictionary information from the European dictionary portal in the field of digital humanities
- Identifying future funding sources and developing collaborative funding applications

There will be a regular exchange between this WG and the other WGs concerning new editorial insights and technical possibilities.

In order to facilitate close collaboration between the WGs, there will be slots for inter-WG meetings at all WG meetings.

Training Schools

In addition, the Management Committee (MC) will organise a series of three Training Schools designed to underpin the activities of WG2, WG3 and WG4 (WG1 will get most of its scientific input from these three WGs). Provisional titles of the Training Schools are as follows:

Training School 1: Standard tools and methods for retro-digitising dictionaries (Year 2)

Training School 2: Innovative approaches in e-lexicography (Year 3)

Training School 3: Lexicography from a pan-European perspective (Year 4)

Workshops

Some of the WG meetings will be preceded by specialist workshops in which experienced Action members and/or visiting experts will share their knowledge in relation to relevant aspects of the theory and methodology of European e-lexicography. These workshops will take place during the whole life of the Action.

Short-Term Scientific Missions

A programme of STSMs will be developed to enable ESRs and other Action members to visit

established dictionary projects and centres of excellence. These visits will range from discussions and demonstrations to possible longer involvement in the activities of the centres visited, in accordance with the needs of the ESRs. They will allow ESRs in particular to build up their own networks and they will facilitate and increase capacity for research in the field of e-lexicography in COST countries. It is envisaged that three STSMs will take place in each year of the Action.

E. ORGANISATION

E.1 Coordination and organisation

This Action will be organised according to “Rules and Procedures for Implementing Cost Actions” 4154/11. The Action will be driven and coordinated by a Management Committee, comprising of an elected Chair and Vice-Chair and up to two representatives from each COST country. The MC will meet every 6 months. The MC will be responsible for co-ordinating the activities of the WGs, budget planning and the allocation of funds, organising Training Schools and conferences, developing and overseeing programmes for STSMs and cascading information from the Action through national and international networks. The MC will monitor progress in relation to the scientific focus and workplan in relation to the achievement of milestones, including electing the Working Groups Leaders; staging Training Schools, launching programmes for STSMs (Year 1), staging an interim conference (Year 2), launching the dictionary portal and the Action website (Year 2), completing other Action deliverables (Year 3 and 4) and staging the final conference (Year 4). The deliverables of the Action will be peer-reviewed.

The intermediate and final conferences as well as two of the WG meetings are planned to be held in connection with relevant conferences in 2015 and 2017; two further WG meetings will be held in connection with relevant conferences in 2014 and 2016.

A Steering Group (SG) will comprise of the Chair and Vice-Chair of the MC, and the Leader and Vice-Leaders of the 4 Working Groups as well as an STSM-manager, a Training School-manager and an ESR researcher-manager. The SG will be responsible for preparing annual reports on the work of the Action, overseeing the development and maintenance of the Action website, and monitoring the COST procedures.

The Action’s website with the European dictionary portal will serve as a central information and dissemination point during the life of the Action and beyond. The website will include an Action calendar and programme, information relating to the Action’s progress and a programme for STSMs and Training Schools. Over the course of the Action, the website will supply resources to facilitate collaborative research. It will act as an intranet to exchange information amongst

participants and an extranet to promote the Action to the wider public.

E.2 Working Groups

The bulk of the coordination work of the Action will be conducted in four Working Groups, which will collaborate closely: WG1 Integrated interface to European dictionary content; WG2 Retro-digitized dictionaries; WG3 Innovative e-dictionaries; WG4 Lexicography and lexicology from a pan-European perspective. Each WG will have a Leader and a Vice-Leader and will produce an annual report on its activities, which will be co-ordinated and summarised by the MC and made publicly available on the Action website.

E.3 Liaison and interaction with other research programmes

The Action will seek cooperation with CLARIN-ERIC (Common Language Resources and Technology Infrastructure - European Research Infrastructure Consortium), which is committed to establish an integrated and interoperable research infrastructure of language resources and their technologies and with DARIAH (Digital Research Infrastructure for the Arts and Humanities), which aims to enhance and support digitally enabled research across the arts and humanities. The Action will invite CLARIN and DARIAH experts to the appropriate WG meetings and will take into account the standards adopted by these two research infrastructures.

E.4 Gender balance and involvement of early-stage researchers

The Action will appoint an ESR/female researcher-manager as part of the SG.

A number of ESRs have been involved in the preparation of this application and the gender balance is well represented among the participants. The MC will encourage ESRs as well as more experienced researchers to increase their skills by participating in Training Schools and Short-Term Scientific Missions. The MC will ensure that ESRs and female researchers share fully in the allocation of responsibilities and activities within these Working Groups and will encourage ESRs and female researchers to become Leaders or Vice-Leaders of Working Groups and to become part of the Management Committee. Furthermore ESRs and female researchers will be given preferential treatment when selecting speakers for Action conferences and workshops.

To support ESRs the Action will cooperate closely with the European Master in Lexicography. It is agreed with the speakers of the EMLex-consortium that the Action will invite Master students to

participate in Working Group meetings and in particular to participate in Training Schools. Further cooperation with the European Master in Lexicography will include developing common topics for PhD theses of Master Students.

An ESR network may be created within the Action as an ESR think tank.

To allow female researchers and other participants with young children to attend WG meetings, conferences etc., important Action events will not take place during main holiday times.

Furthermore the Action will provide childcare during Action events as necessary.

The applications for this COST Action reflect a wide European dimension, including participants from all parts of Europe. The MC will also ensure that the Leaders and Vice-Leaders of the Working Groups represent a wide range of European countries. Furthermore the Action participants reflect a wide range of dictionary types: synchronic – diachronic/historical, standard language – minority language and small language dictionaries.

F. TIMETABLE

	Year 1		Year 2	
	Semester 1	Semester 2	Semester 1	Semester 2
MC meeting	X	X	X	X
WG meetings	--	X	X	X
STSM	--	3 STSMs	3 STSMs	3 STSMs
Training Schools	--	--	X	--
Conferences	--	--	--	X
	Year 3		Year 4	
	Semester 1	Semester 2	Semester 1	Semester 2
MC meeting	X	X	X	X
WG meetings	X	X	X	--
STSM	3 STSM	3 STSMs	3 STSMs	1 STSM
Training Schools	X	--	X	--
Conferences	--	--	--	X

The two conferences and four of the Working Group meetings will be held in connection with relevant conferences.

G. ECONOMIC DIMENSION

The following COST countries have actively participated in the preparation of the Action or otherwise indicated their interest: CH, DE, DK, ES, FI, FR, HU, IT, LV, NL, NO, PL, PT, RO, SE, UK. On the basis of national estimates, the economic dimension of the activities to be carried out under the Action has been estimated at 64 Million € for the total duration of the Action. This estimate is valid under the assumption that all the countries mentioned above but no other countries will participate in the Action. Any departure from this will change the total cost accordingly.

H. DISSEMINATION PLAN

H.1 Who?

The target audiences for the dissemination of the results of the Action include both academic and commercial lexicographers, linguists and scientists from neighbouring disciplines from all over Europe and worldwide. The European dictionary portal will disseminate lexical information on the languages of Europe to the general public. The Action will be contacting policy- and decision-makers at the regional, national and European level (Ministries of Education, of cultural heritage etc.) to increase the opportunities for the methodologies and standards developed by the Action to be implemented.

H.2 What?

The dissemination of the results of the Action will take place on various levels:

- a) The European dictionary portal will make lexical information on the languages of Europe accessible to the general public, to lexicographers and to other scientists. The results obtained in the WGs will also feed into the presentation of information in the dictionary portal.
- b) The public part of the Action website will provide the annual reports of the WGs, the roadmaps developed in the Action, a calendar detailing current and upcoming Action events and other documents created within or relevant to the Action.
- c) The password-protected part of the Action website will be used to publish draft or working documents and, if necessary, to create discussion forums for the WGs.
- d) Social media accounts (Twitter, Facebook) will be set up and monitored by the MC in order to disseminate updates to Action members and to other stakeholders.
- e) Members of the Action will submit joint papers to reviewed journals in the fields of lexicography/lexicology, corpus linguistics and computer linguistics in order to gain a broad

academic audience.

f) Email invitations to the kick-off meeting and the interim conference (year 2) will be sent to lexicographers and computational linguists who might be interested in participating in the Action.

g) The final conference (year 4), which will be held in connection with a large conference on (electronic) lexicography, will disseminate the end results of the Action across all WGs and to other stakeholders through the conference proceedings.

h) Training Schools will be held to update ESRs and other Action members on to the state of the art with respect to the topics treated in WGs 2 to 4 and to introduce them to new findings and methods.

i) Other potential stakeholders will be invited to participate in WG meetings.

H.3 How?

a) The Action website with its public and password-protected areas will constitute the main dissemination platform. It will be used to store all outputs and resources generated by and relevant to the COST Action.

b) The European dictionary portal will link the Action to the general public and non-expert dictionary users. It will provide authoritative lexical information on the languages of Europe.

c) Social Media and email will be used to distribute information, alert Action members to new content on the website and generate discussions between workshops and meetings.

d) Action participants will engage in dissemination activities within their own region and professional setting (for example, lectures, courses, workshops).

The dissemination activities will be regularly monitored by the MC and the dissemination strategy updated as necessary. The Action will strongly promote open access and will guarantee that for the Action results; for the dictionary data and tools, policy is determined by each member's organisation.