



EUROPEAN ASSOCIATION FOR LEXICOGRAPHY

EURALEX NEWSLETTER

Editor: Henri Béjoint. Address for correspondence and contributions:
EURALEX Secretary
P.O. Box 1017, DK-1007, Copenhagen K, Denmark

The EURALEX Newsletter

This quarterly Newsletter is intended to include not only official announcements but also news on EURALEX members, their publications, career moves and (it is hoped) their opinions. Please try to support this by sending newsletter contributions to the Secretary at the above address. The deadlines for winter, spring, summer and autumn issues respectively are 15 September, 15 December, 15 March and 15 June annually.

EC Programmes: Linguistic Research and Engineering 1

Since 1992/93, the European Commission (DG XIII, Luxembourg) has been running a research and development programme on *Linguistic Research and Engineering* (LRE). Within this framework, a number of activities are going on which are either of potential relevance to lexicography, directly involve lexicographers or are supposed to lead to fragments of dictionaries, grammars, to textual corpora, annotation schemes, tools for lexical processing and corpus exploration.

It may be useful to include in this newsletter a few reports on such activities, to inform the EURALEX members about projects, work groups, tools, data or simply contact points of potential interest.

We begin a series of short summaries of such activities with a description of EAGLES and more precisely of a work group within EAGLES dealing with lexical description.

EAGLES

EAGLES (= Expert Advisory Group on Linguistic Engineering Standards) is a working group initiated and supported by the European Commission, under the LRE programme. EAGLES was launched in February 1993. Its objective is to pave the way for the construction of commonly accepted methods, tools and guidelines for descriptive and processing work in Speech and Natural Language Processing. It is assumed that on the one hand common, shareable descriptions, methods and tools will have a positive impact on collaboration and exchange of ideas in the field of language engineering, and, on the other hand, that such a common view and shared approaches cannot be in any way imposed, but only emerge as the result of a consensus building process.

EAGLES has five thematically organized work groups, one of which deals with computational lexicons. The other groups concern textual corpora, spoken language and speech processing, computational formalisms for Natural Language Processing and the evaluation of computational tools and components.

The lexicon group of EAGLES has so far concentrated on the following topics:

- o methods for the defining guidelines and standards for computational lexicons;
- o general guidelines for the architecture of computational lexicons (organization, types of information, representation);
- o morphosyntactic description for computational lexicons.

Work on morphosyntax has led to concrete results: the work done so far in a number of projects aiming at building dictionaries (or dictionary models) for computational treatment has been carefully surveyed, in terms of the linguistic distinctions made for each word class (for nouns, for example, not only gender, number and case are distinguished, but some dictionaries also distinguish between common nouns and proper nouns, some have codes to distinguish “count nouns” from “mass nouns”, or they give indications as to the inflection type). The EAGLES survey includes as well the distinctions made in the major “tagsets” for linguistic corpus annotation, since EAGLES assumes that morphosyntactic dictionaries and annotation schemes for textual corpora have a common conceptual basis, although corpus annotation is often a specific application of a more general lexical description and thus introduces additional features not present or not needed in dictionaries.

The morphosyntactic survey is based on a large number of lexicon and corpus encoding schemes. These have been compared, and the minimal consensus between these has been determined, in terms of types of linguistic information recorded. From this comparison, a set of recommendations has been developed. These recommendations fall into an obligatory part (part-of-speech classification), a recommended part (morphosyntactic properties relevant for describing agreement phenomena) and optional extensions specific of a given phenomenon, language or application.

The EAGLES proposals have been developed by the members of the lexicon work group in constant interaction with the corpus work group (members of both come from industrial and academic research centres). Detailed proposals for an application to seven out of the nine languages of the European Community have been made, all following the basic EAGLES proposals and again comparing these with the most prominent lexical and corpus description schemes in the individual countries (for example INL and CELEX for Dutch).

EAGLES is in the process of documenting the results of its first phase of work, and from October 1994 on, reports on the work of all EAGLES groups will be available. The lexicon group will then deepen its work on morphosyntax, by preparing a set of guidelines for the application of the proposed morphosyntactic classifications, including a description of the criteria applied to decide about how to actually classify a given word form. Giving such criteria is in many cases very difficult; for example, what are good tests for separating count noun meanings from mass noun meanings of a given item? The lexicon group would thus welcome comments, feedback and critical reactions to its proposals, in particular from lexicographers.

The work group also aims at preparing lexical lists encoded according to the proposed guidelines, and at investigating ways and means of mapping existing descriptions onto the EAGLES proposal.

Additional topics which the EAGLES lexicon group is working on include the description of syntactic subcategorization (transitive, intransitive, etc.) and contrastive description for bilingual and multilingual dictionaries.

The president of the EAGLES lexicon work group is Dr. Annie Zaenen,

Rank Xerox Research Centre

Immeuble le Quartz

6, chemin de Maupertuis

F-38240 Meylan

France

Tel: +33-76-61-50-73
 Fax: +33-76-61-50-99.

The group is hosted by GSI-ERLI c/o Dr. Simon Sabbagh
 GSI ERLI
 1, place des Marseillais
 F-94227 Charenton-le-Pont Cedex
 France

Tel: +33-1-48.93.81.21
 Fax: +33-1-43.75.79.79.

Reports can be received as well from the internal editor, Ulrich Heid
 Institut für Maschinelle Sprachverarbeitung-Computerlinguistik
 Azenbergstrasse 12,
 D-70174 Stuttgart
 Germany
 Tel: +49-711-121-1373
 Fax: +49-711-121-1366.

Lexicography in Bulgaria

The Department of Lexicology and Lexicography of the Institute for the Bulgarian Language of the Bulgarian Academy of Sciences is preparing an extensive *Dictionary of the Bulgarian Language*. Volume XI is now finished, and there are seven or eight more volumes to come. The Department has 18 lexicographers and eight linguists working on the project.

Dr. Kassabov has recently published a book on the theory and practice of semantics in dictionaries. Dimitr Popov's *Monolingual Bulgarian Dictionary* – a desk-sized dictionary of the contemporary Bulgarian language – will be published soon.

Additional information on the activities of the group can be obtained from Sabina Pavlova, ul. "Shipchenski prohod" 52, bl. 17, Institute for the Bulgarian Language – BAN, Sofia 1113, Bulgaria.

New publications

Corréard M.H. and V. Grundy, 1994, *The Oxford – Hachette French-English/English-French Dictionary*, Oxford University Press, Oxford, UK, and Hachette, Paris, France.
 Dubichinsky, V.V., 1993, *Lexical Parallels*, Kharkov.

Dubichinsky, V.V., S.V. Vakulenko and A.I. Danilenko (eds), 1992, *Current Problems of Lexicography*, Kharkov.

Dubichinsky, V.V., V.A. Zubareva and S.P. Oleynik, 1992, *Russian-English Explanatory Dictionary of Collocational Terms from Foreign Trade*, Kharkov.

Lysyenko, L.A. and S.I. Doroshenko (eds), 1992, *Linguistic Studies*, Kharkov.

Strauss, E., 1994, *Dictionary of European Proverbs*, Routledge, London.

Forthcoming events

1994

17–24 October Buenos Aires, Argentina: 4th Ibero-American Terminology Seminar.
 Info: SHT, Villa Ocampo, Bloriondo 1811, 1643 Beccar, Argentina OR Union Latine,
 14 Bd Arago, F-75013 Paris, France. Tel: 33 1 43 36 14 14. Fax: 33 1 45 35 75 01.

24–25 November Vienna, Austria: 3rd TermNet Symposium "Terminology in Advanced

Microcomputer Applications". Info: TermNet, Grüng. 9/17, A-1050 Vienna, Austria. Fax: 56 77 64.

27–30 December New York, USA: Conference of the Modern Language Association of America. Info: MLA, 10 Astor Place, New York 10003–6981, USA.

1995

31 March–3 April York, UK: Conference of the Association for Language Learning (ALL), "Language world". Info: ALL, C. Wilding, 16 Regent Place, Rugby CV21 2PN, UK.

19–22 April Maastricht, The Netherlands: Maastricht session of the 2nd international Maastricht – Łódź Duo Colloquium on "Translation and Meaning" (see 1995, 22–24 September for the Łódź session). Info: Dr. Marcel Thelen, Rijkshogeschool Maastricht, Faculty of Translation and Interpreting, P.O. Box 964, 6200 AZ Maastricht The Netherlands. Tel: 31 43 46 66 40, 31 43 46 64 71 (direct line). Fax: 31 43 46 66 49.

May Alushta, Crimea, Ukraine: 2nd International Conference on Lexicography. Info: Vladimir Dubichinsky, Kharkov Lexicography Society, Kharkov Polytechnical University, 5a Vesnina St., Kharkov, Ukraine 310023.

9–11 June Copenhagen, Denmark: 3rd Language International Conference on Teaching Translation and Interpreting. Info: Cay Dollerup, Centre for Translation and Lexicography, University of Copenhagen, 96 Njalsgade, DK 2300 Copenhagen S. Tel: 45 31 54 22 11. Fax: 45 32 96 15 18.

11–13 July Luxembourg: MT SUMMIT V, Fifth Congress of the International Association of Machine Translation.

19–24 July Beijing, China: 22nd International Systemic Functional Conference. Info: Hu Zhuanglin, Department of English, University of Beijing, Beijing, 100871, China.

29 August–2 September, Vienna, Austria: 10th European Symposium on Languages for Special Purposes "Multilingualism in Specialist Communication". Info: Gerhard Budin, University of Vienna, Infoterm. Tel: 43 1 26 75 35 310. Fax: 43 1 216 32 72.

22–24 September Łódź, Poland: Łódź session of the 2nd international Maastricht – Łódź Duo Colloquium on "Translation and Meaning" (see 1995, 19–22 April for the Maastricht session). Info: Dr. Habil. Barbara Lewandowska-Tomaszczyk, Institute of English Studies, University of Łódź, Al. Kosciuszki 65, 90–514 Łódź, Poland. Tel: 48 42 36 63 37. Fax: 48 42 36 63 37, 48 42 36 68 72. Telex: 88 62 91 UL PL. e-mail: ZAJEANFA at plunlo51.bitnet.

Late September Lyon, France: Fourth Colloquium of AUPELF-UREF, (Lexicology, Terminology and Translation network) "Lexicomatique et dictionnaires". Info: P. Thoiron, CRTT, Université Lumière Lyon 2, 86 rue Pasteur 69365 Lyon Cedex 07, Tel: 78 69 72 13. Fax: 78 69 56 01. e-mail; thoiron@diogene.univ-lyon2.fr.

27–30 December New York, USA: Conference of the Modern Language Association of America. Info: MLA, 10 Astor Place, New York 10003–6981, USA.

1996

29 March–1 April Exeter, UK: Conference of the Association for Language Learning (ALL), "Language world". Info: ALL, C. Wilding, 16 Regent Place, Rugby CV21 2PN, UK.

15–19 July Sidney, Australia: 23rd International Systemic Functional Congress. Info:

Diana Slade, Centre for Language and Literacy, School of Adult and Language Education, Faculty of Education, Sidney University of Technology, PO Box 123, Broadway 2007, New Wouth Wales, Australia. e-mail: D. Slade at UTS.EDU.AU.

4–9 August Jyväskylä, Finland: (11th) World Congress of the International Association of Applied Linguistics (AILA). Info: Prof. Kari Sajavaara, Department of English, University of Jyväskylä, SF-40100 Jyväskylä, Finland. Tel: 358 41 601221. Fax: 358 41 601213.