

SPRING 1998

# EURALEX NEWSLETTER



Editors: Carla Mareello and Rosamund Moon.

*Address for correspondence and contributions:* Carla Mareello, Corso Unione Sovietica 115, I-10134 Torino, Italy. E-mail: [mareello@cisi.unito.it](mailto:mareello@cisi.unito.it) and [rosamund@cobuild.collins.co.uk](mailto:rosamund@cobuild.collins.co.uk)

---

## 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 Carla Mareello at the above address. The deadlines for winter (December), spring (March), summer (June), and autumn (September) issues are respectively 15 September, 15 December, 15 March, and 15 June annually.

## The EURALEX Web Site

The URL of the EURALEX web site is <http://www.ims.uni-stuttgart.de/euralex>.

## IJL e-mail

The new e-mail address for Tony Cowie, editor of IJL, is A.P. Cowie@btinternet.com.

## Helmut Feldweg

It is with deep and profound regret that we find ourselves in the position of announcing the unexpected and sudden death of our colleague, Helmut Feldweg. Helmut, at the age of 41, succumbed to a previously undetected cancer of the heart this past Monday, November 24th. For many of us, Helmut was more than just another colleague at the University of Tübingen: he was a friend, a partner, a solid rock in the whirls and eddies of research and life, and the one we went to when we had any questions.

After studying Chinese at the University of Göttingen in the early eighties, Helmut worked at the Max-Planck-Institute for Psycholinguistics in Nijmegen,

where he was involved with database programs such as CELEX and CHILDES. At the University of Tübingen, where he has been since 1992, Helmut developed a Part-of-Speech tagger for German (called LIKELY), worked on the development of lexical knowledge bases, and in cooperation with several European partners, established a prototype for an intelligent, context-sensitive on-line dictionary look-up system. At the time of his death, he was engaged in building up a German counterpart to the semantic on-line lexical reference system WordNet (developed at Princeton).

Interacting with Helmut, whether it was at project meetings, over lunch, on a hike, or while flying kites, was always a singularly positive experience, which can be traced directly to the solid competence and imagination that Helmut combined with the utter absence of a need to put himself before others. So, for example, Helmut (without ever seeking or needing an acknowledgement of this fact) quietly but effectively trained, pushed, and supported a good number of the women who have to date advanced from the Institute of Linguistics at the University of Tübingen to other jobs in both industry and academia.

In short, Helmut Feldweg was a pleasure and an inspiration to work with, and the brevity of his research career is all the more to be regretted. We will miss all aspects of his personality, but perhaps most devastatingly, we will miss the good humour and dry wit with which he confronted both linguistics and life, and which made him such a wonderful colleague.

*Miriam Butt, FG Sprachwissenschaft, Universität Konstanz (formerly at Tübingen)*

*Arnim von Stechow, Seminar für Sprachwissenschaft, Universität Tübingen*

## News from Exeter and the Dictionary Research Centre

Reinhard Hartmann and his co-author Gregory James (Hong Kong University of Science and Technology) have completed their 'Dictionary of Lexicography', which will be published by Routledge in March 1998. The dictionary gives a comprehensive overview of the current state of lexicography and its interdisciplinary contacts within a wider context of 'reference science'. The representative literature has been included and an alphabetically-arranged appendix lists all bibliographical references given in the more than 2,000 entries, which also provide examples of relevant reference works.

The new intake of five full-time and one part-time students on the M.A. in Lexicography are settling down to their course options. The topics of M.A. dissertations and Ph.D. theses in progress are listed in Newsletter 15 which has recently been dispatched from the Dictionary Research Centre. This also contains a report on the 1997 InterLex Course, which brought 20 participants from 9 countries to Exeter. InterLex 12 will be held from 27 April to 1 May 1998.

*Reinhard Hartmann, Exeter*

## Asialex

The second newsletter is now ready for distribution, and it will be sent to all Asialex members. Non-members who are interested in getting a copy should e-mail or write to Amy Chi. The latest issue includes an introduction from the

Asialex President and the latest developments in the Association. For example, ASIALEX is planning to hold a conference on bilingual lexicography (tentatively) around December of 1998 in Guangzhou, China.

*Amy Chi* (e-mail [lcamychi@usthk.ust.hk](mailto:lcamychi@usthk.ust.hk))

*c/o Language Centre, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

## The SALEX course, Grahamstown, September 1997

A two-week course in lexicography was held at Rhodes University in South Africa in September 1997. Under the new South African constitution, there are now 11 official languages, and plans are afoot to create state-funded Dictionary Units for the 9 African languages, to complement the already-existing units for English and Afrikaans. Against this background, and under the auspices of Afrilex, Penny Silva – Director of the Dictionary Unit for South African English – set about organizing a foundation course to train the next generation of South African lexicographers.

Three UK lexicographers, Sue Atkins (course leader), Michael Rundell, and Edmund Weiner, were invited to design and teach this course, which became known as SALEX (“South African lexicography”). The challenge was to produce an intensive two-week programme that would be language-independent, practically-oriented, and firmly situated in the empirical and descriptive tradition. The objective was to give the participants the basics of hands-on lexicography, in order to equip even those with no dictionary-making experience to set up a Dictionary Unit, plan a project, recruit and train a team, design a dictionary, and see it through to publication.

Thirty-five linguists, representing over a dozen languages, attended the course, which gave an overview of every aspect of dictionary production, from project management and corpus development to the actual compilation and editing of dictionary text. Topics covered included project planning and scheduling, style guides, corpus design criteria, and the nuts and bolts of writing dictionary entries. Key concepts from a range of relevant academic disciplines (including lexical semantics, language-acquisition theory, and computational linguistics) were also discussed, and teaching took the form of lecture sessions alternating with practical workshops.

A key feature of the course was its exhaustive documentation: all of the material was saved in MS Powerpoint files, and this enabled us to give participants complete copies of overheads and worksheets at the end of each session. By the end of the course, they took away with them a large file of material which they could later use to train members of their own Units.

SALEX received generous support from the British Council, and also from OUP, CUP, Collins Cobuild, and Addison Wesley Longman, who all donated dictionaries. In addition corpus text and software tools, loaned by Longman Dictionaries, gave participants hands-on experience of language-analysis and dictionary-compilation programs. Two South African corporate donors, First National Bank and The Anglo-American and De Beers Chairman’s Fund, also helped to make the course possible. A follow-up course is now being planned

for 1998, and it will look in more detail at bilingual lexicography and focus in greater depth on project management issues such as recruitment and scheduling.

Course organizer, Penny Silva, said afterwards that SALEX had fulfilled all her expectations, and one of the participants, Prof Sizwe Satyo of the University of Cape Town, sent her this message: 'Not only did we learn about dictionary-making as such, our lecturers virtually took us on a guided tour around our own brains to help us observe just how language is computed. It was an unforgettable experience'.

For more information, including the full programme, visit the SALEX web site: <http://www.ru.ac.za/affiliates/dsae/salex97/>

*Michael Rundell*

### Workshop: SENSEVAL and the Lexicography Loop

This workshop is sponsored by EURALEX and by the Association for Computational Linguistics (Lexicons Special Interest Group).

One of the core tasks of lexicography is determining what senses a word has. This task is intimately related to another, of determining, for an instance of a word in context, which of the word's senses applies. The task is often called 'word sense disambiguation' (WSD). While for people, WSD seems to be instant and effortless, for computers it is very hard. Yet computers need to do WSD for, among other things, machine translation. It is required for selecting the correct translation wherever different senses of source-text words have different translations: also for lexical access, as online dictionaries will be more useful if they take you to the correct sense of a word-in-context, not merely the correct headword.

There are now many automatic WSD programs but it is currently very hard to determine which are better, which worse, and where the strengths and weaknesses of each lie. There is widespread agreement that the field urgently needs an evaluation framework. Under the auspices of EURALEX and ACL (SIGLEX), we are setting up a pilot evaluation exercise for 1998. We shall be undertaking evaluation for at least English, French, Italian and Spanish. The performance of WSD systems, and other issues relating to the task, will be discussed at a workshop to be held at: Herstmonceux Castle, Sussex, UK, from 2-4 September 1998.

Lexicography is central to the enterprise. Most WSD programs make use of online dictionaries, and the quality and style of the dictionary is then crucial. Also the preparation of a set of correct answers, against which the programs will be evaluated, is a demanding lexicographic task.

If you have skills or resources that you would like to contribute to the exercise, or if you have a working WSD program (or will have one by Summer 1998), and would like to subject it to objective, quantitative evaluation, please mail your expression of interest to: [senseval-coord@itri.bton.ac.uk](mailto:senseval-coord@itri.bton.ac.uk).

For more information, see the SENSEVAL web site at <http://www.itri.bton.ac.uk/events/senseval/euralex-cfp.txt>.

*Adam Kilgarriff, University of Brighton*

## Lexicographical Work in Progress

### COMLEX and Proteus Project

The Proteus Project at New York University would like to announce that the new version of COMLEX Syntax will include classification for adverbs. (COMLEX Syntax is a broad-coverage English lexicon of syntactic features designed for natural language processing by computer. This dictionary was developed at New York University and is distributed by the Linguistic Data Consortium. See our web site at <http://cs.nyu.edu/cs/projects/proteus/comlex> for more information.) The adverbs will be classified as to syntactic position, i.e. pre-adjective (pre-adj), post-noun etc, and will be given features such as MANNER-ADV and TEMPORAL-ADV as well as some META-features e.g. attitude and epistemic. The new version of COMLEX Syntax, which includes the adverbs as well as some corrections of the original dictionary, should be available from the Linguistic Data Consortium in January, 1998.

*Catherine Macleod, Ralph Grishman and Adam Meyers, New York*

### A Persian–English Dictionary

This dictionary is intended to be a tool for translating Neopersian lyrical poems into English. The project began in October 1995, and is expected to continue until December 1999: it is being funded by the University of Venice, and the Ministry of University and Scientific-Technological Research. The project team includes Riccardo Zipoli, Daniela Meneghini, Valentina Zanolla, and Hamid Reza Baharlu. An account of the project is given in *Eurasiatica 51* (Venice, 1997): D. Meneghini Correale, V. Zanolla, and R. Zipoli, 'Outline of a Persian-English Dictionary'.

The dictionary will be unidirectional, and will deal with the language of the Classical form of the ghazal, from the 12th century to the 19th century. It will be accompanied by a database of the corpus which underpins the compilation of the dictionary: a corpus consisting of, at least, thirty samples of 1,000 lines, chosen from the collections of the most important Persian poets. Examples in the dictionary will be taken from this corpus. The dictionary is expected to consist of 18–20,000 entries, and it will include all the lemmas found in the corpus, together with the irregular verbal stems, broken plurals, plurals with a meaning not directly deducible from the singular, and graphic-phonetic variants found in the corpus. Each entry will consist of the headword in Arabic-Persian characters, a transliteration, cross-references as appropriate, semantic equivalent or equivalents, contextual or register information, a verse citation (with English translation) to illustrate the use of the headword, grammatical notes, and phrases accompanied by verse citations with translation, and statistical data (frequencies, Chi square, and usage coefficient of the headword). The dictionary will be printed but it will be also distributed on CD-Rom.

*Dr. Daniela Meneghini, University of Venice, Dipartimento di Studi Eurasiatici, San Polo 2035, Venezia 30125, Italy*

## A French–English Bilingual Dictionary of Retailing/Dictionnaire bilingue de la distribution

This dictionary is due to appear in Spring 1998. It is, to the compilers' knowledge, the first (French–English) bilingual dictionary entirely and uniquely devoted to retailing. To fill this gap and satisfy the needs of many translators and terminologists, as well as the linguistics/translation services of retailers across English and French speaking countries, we are launching a semi-encyclopedic bilingual dictionary focusing on the 500 basic concepts of retailing. These 500 dictionary entries enable us to refer to some 4,000 or 4,500 terms in English and in French – which, according to our research, will give a fairly comprehensive view of this specialized field. The dictionary will be published on paper, but also as an electronic dictionary (based on hypertext techniques).

*Jeanne Dancette, Scientific Director & Project Manager*

(514) 343–7456; e-mail: dancettj@ere.umontreal.ca

*Christophe Rethore, External Affairs & Public Relations*

(514) 343–7866; e-mail: rethorec@ere.umontreal.ca

## TELRI – East meets West: A compendium of multilingual resources

The TELRI Association announces the publication of a double CD-Rom, containing extensive corpora, both spoken and written, in more than 20 languages of Western, Central and Eastern Europe. The resources are available in plain text, SGML and HTML, and have largely been successfully aligned. Also available are tools including a corpus query language, concordancer, alignment tools, software, taggers, and samples of research work involving the data.

The CD-ROM package is usable on UNIX, PC, and Mac platforms and is available in a limited edition for academic purposes only at a non-profit cost of 25 ECU. To request further information or a copy of the product, please contact [telri@ids-mannheim.de](mailto:telri@ids-mannheim.de). Please see also:

<http://www.ids-mannheim.de/telri/cdrom.html>.

*Dr Ann Lawson, Abteilung LEXIK, Institut für deutsche Sprache*

## Forthcoming events

### 1998

#### April

1–4, Stuttgart, Germany: 3rd International Symposium on Phraseology. Info: 3rd International Symposium on Phraseology, c/o Dr. Ulrich Heid, Universität Stuttgart, IMS-Computerlinguistik, Azenbergstr. 12, D 70174 Stuttgart, Germany. Telefax: +49 711 121 1366. E-mail: [isp-3@ims.uni-stuttgart.de](mailto:isp-3@ims.uni-stuttgart.de)

16–19, Pennsylvania, USA: LSRL 28 – XXVIII Linguistic Symposium on Romance Languages. Info: LSRL 28 Organizing Committee (Marc Authier, Barbara Bullock, & Lise Reed), Department of French, The Pennsylvania State University, Room 325 – Borrowes Building, University Park, PA

16802–6203, USA. Tel: +1 814 863–2814. E-mail: LSRL@psu.edu. Web site: <http://www.psu.edu/lsrc/>

24–25, Köln, Germany: German Terminology in International Business, 6th Symposium of the German Terminology Society (DTT). Info: Jens Hiltmann, c/o Translingua, Ludwig-Erhard-Allee 3, D-53175 Bonn, Germany. Tel: +49 228 8160 152. Fax: +49 228 30888 55. E-mail: [jensh@translingua.de](mailto:jensh@translingua.de)

27–1 May: 12th International Course on Lexicography. Info: Dictionary Research Centre, University of Exeter, EX4 4QH. Fax +44 1392 264361. E-mail [r.r.k.hartmann@exeter.ac.uk](mailto:r.r.k.hartmann@exeter.ac.uk)

#### May

5–9, Évora, Portugal: SILF98 – XXII International Congress of Functional Linguistics. Info: Comissão Organizadora do XXII Colóquio Internacional de Linguística Funcional, Departamento de Linguística e Literaturas, Universidade de Évora, 7001 Évora Codex, Portugal. Tel: +351 66 744675. Fax: +351 66742597. E-mail: [SILF98@evunix.uevora.pt](mailto:SILF98@evunix.uevora.pt)

6–8, Barcelona, Spain: 4th International Conference on Translation and Interpretation. Info: Secretaria del Departament de Traducció i d'Interpretació, Facultat de Traducció i d'Interpretació (Edifici M-1), Campus Universitari s/n, 01893 Bellaterra, Barcelona, Spain. Tel: +34 93 581 2761. Fax: +34 93 581 2762. E-mail: [congres\\_ti@cc.uab.es](mailto:congres_ti@cc.uab.es) Web site: [http://cc.uab.es/congres\\_ti/](http://cc.uab.es/congres_ti/)

1–6, Jyväskylä, Finland: Summer School of Applied Linguistics. Info: Prof. Kari Sajavaara, Centre for Applied Language Studies, University of Jyväskylä, P.O. Box 35, 40351 Jyväskylä, Finland. Tel: +358 14 603 529. E-mail: [sajavaara@cc.jyu.fi](mailto:sajavaara@cc.jyu.fi)

24–26, Regensburg, Germany: International Symposium, Degrees of Restructuring in Creole Languages. Info: Prof. Dr. Ingrid Neumann-Holzschuh, Regensburg Universität, Institut für Romanistik, D-93040 Regensburg, Germany. E-mail: [ingrid.neumann-holzschuh@sprachlit.uni-regensburg.de](mailto:ingrid.neumann-holzschuh@sprachlit.uni-regensburg.de)

#### July

17, Brisbane, Australia: AUSTRALEX biennial conference. Web site: <http://www.anu.edu.au/linguistics/alex>

23–29, Bruxelles, Belgium: XXIIe Congrès International de Linguistique et Philologie Romanes. Info: XXIIe Congrès International de Linguistique et Philologie Romanes, CP 175/Faculté de Philosophie et Lettres, Université Libre de Bruxelles, 50 avenue Franklin D. Roosevelt, B-1050 Bruxelles, Belgium. Tel: +32 2 650 24 36. Fax: +32 2 384 04 83. E-mail: [congres@romane.ulb.ac.be](mailto:congres@romane.ulb.ac.be). Web site: <http://www.emich.edu/~linguist/issues/8/8-124.html#1>

#### August

4–8, Liège, Belgium: EURALEX '98. Info: Congress Organizers EURALEX '98, University of Liège, Department of English Language and Linguistics, Building A2, 3, Place Cockerill, B-4000 Liège, Belgium. Tel: +32 4 3665360 (Michiels). Fax: +32 4 3665721. E-mail: [amichiels@ulg.ac.be](mailto:amichiels@ulg.ac.be). Web site: <http://engdepl.philo.ulg.ac.be/euralex.htm>

10–14, Montréal, Canada: COLING/ACL-98, 17th International Conference on Computational Linguistics and 36th Annual Meeting of the

Association for Computational Linguistics. E-mail: [coling-acl98@iro.umontreal.ca](mailto:coling-acl98@iro.umontreal.ca). Web site: <http://coling-ac198.iro.umontreal.ca>

17–28, Saarbrücken, Germany: 10th European Summer School in Logic, Language and Information. Includes a one-week course on ‘Word Sense Disambiguation’, and other workshops and courses related to the lexicon. Info: DFKI, ESSLLI X, Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany. Phone: +49 681 302 4933. Fax: +49 681 302 4929. E-mail: [esslli@coli.uni-sb.de](mailto:esslli@coli.uni-sb.de). Web site: <http://www.coli.uni-sb.de/esslli>

26–30, St. Andrews, Scotland: 31st Annual Meeting of the Societas Linguistica Europaea. Info: Dr. C. Beedham, Dept of German, School of Modern Languages, The University, Buchanan Bldg., St. Andrews, Fife KY16 9PH, Scotland. E-mail: [cb1@st-andrews.ac.uk](mailto:cb1@st-andrews.ac.uk)

September

2–4, Herstmonceux Castle, Sussex, England: Workshop, ‘Word Sense Disambiguation Evaluation Exercise’, sponsored by EURALEX/ACL. (See the article above.) Info: Adam Kilgarriff, SENSEVAL, ITRI, University of Brighton, Lewes Road, Brighton BN2 4GJ, England. E-mail: [senseval-coord@itri.bton.ac.uk](mailto:senseval-coord@itri.bton.ac.uk). Web site: <http://www.itri.bton.ac.uk/events/senseval/euralex-cfp.txt>

17–19, Reading, England: The Convergence and Divergence of Dialects in Changing Europe. Info: Paul Kerswill (ESF), Dept. of Linguistic Science, The University of Reading, Whiteknights, P.O. Box 218, Reading RG6 6AA, England. E-mail: [p.e.kerswill@reading.ac.uk](mailto:p.e.kerswill@reading.ac.uk)

23–26, Granada, Spain: 2nd International Congress of the European Society for Translation Studies. Info: EST Congress – Granada 1998, Organizing Committee, Gymnasiumstrasse 50, A-1190 Vienna, Austria. Web site: <http://www.univie.ac.at/transvienna/est/>

October

2–3, Savonlinna, Finland: Interdisciplinary Colloquium on Translation and Cognition. Contact: Riitta Jääskeläinen, University of Joensuu, Savonlinna School of Translation Studies, P.O. Box 48, 57101 Savonlinna, Finland. E-mail: [riitta.jaaskelainen@joensuu.fi](mailto:riitta.jaaskelainen@joensuu.fi)

8–10, Milano, Italy: IX International Congress of Linguists. Info: Prof. Giancarlo Bolognesi, Istituto di Glottologia, Università Cattolica, Largo Gemelli 1, 20123 Milano, Italy. E-mail: [istglott@mi.unicatt.it](mailto:istglott@mi.unicatt.it)

November

16–21, Habana, Cuba: IV IberoAmerican Symposium of Terminology. Organizing association: RITerm-Red Iberoamericana de Terminología (IberoAmerican Terminology Network). Info: Manuel Barreiro Sánchez, Comité Nacional Preparatorio del VI RITerm, 19 de mayo, no 14, esq. Ayestarán – Municipio Plaza, Ciudad de la Habana, Cuba CP 10600. E-mail: [ulcuba@ceniai.inf.cu](mailto:ulcuba@ceniai.inf.cu)