

elsnews

The Newsletter of the European Network in Language and Speech

Dec 1996

ELSNET-2:

Preparing for the next Millennium.

Steven Krauwer, Utrecht University

ELSNET has just entered a new phase: the EC contract that funded our activities over the last four years expired on October 31st, and we have made a seamless transition to a new contract, which will finish at the end of 1999.

The new contract does not imply that we have become a new network: ELSNET, the organism established five years ago, continues to grow, and will thrive well into the future. Another recent change which supports this continuity is the establishment of the ELSNET Foundation, a legal entity under Dutch law. This change in legal status has a number of advantages. First of all, the ELSNET Foundation provides ELSNET with an identity which is independent of potential discontinuities of funding. Secondly, it can act on behalf of ELSNET whenever a legal entity is necessary or desirable. And finally, it makes it possible for the network to propose and coordinate projects, or to join projects as a full partner. Although it is clearly not ELSNET's task to carry out research, the creation of the ELSNET Foundation means that it can now participate in research with a view to training, information exchange, or other support activities.

ELSNET's working structures will basically remain unchanged, except for the creation of an Industrial Panel, which will assist the Executive Board in determining its long term goals and priorities. In terms of research activities, ELSNET-2 will focus on the (further) integration of language and speech, the integration and comparative evaluation of research results, the relationship between academia and industry, and language coverage. The Network Programme, which can be found on our web pages, gives a detailed description of these.

ELSNET's policy of extending itself towards Central and Eastern Europe, through its Copernicus project Elnet-Goes-East, will be continued. Additional actions will be organized to establish connections with academic and industrial sites in the Mediterranean area. By the year 2000 the result should be a truly Pan-European Network of Language and Speech, covering all of Europe and its closest neighbours.

What else does the future hold in store? The results of an EC review of the concept of Networks of Excellence have recently become available, and these are obviously of great relevance to us. The panel's general conclusion is that Networks of Excellence constitute a sound and cost-effective way of structuring and coordinating R&D and related activities in the technological domain. They have been successful at promoting knowledge and information exchange, and can consolidate or increase Europe's lead in areas of commercial or technological strength.

But the panel also draws attention to a few areas where it feels there is room for greater effectiveness and potential impact. One area of concern is the existence of so-called 'passive' nodes. The panel strongly recommends that networks take measures to increase the involvement of such nodes. As I have said before, a network is about sharing, not only of knowledge and expertise, but also of problems and needs. As far as ELSNET is concerned, we intend to contract out more tasks as small projects to members (instead of leaving their execution to members of standing committees). We are also planning a Network conference, to take place in the second half of 1997, as an opportunity for a wide range of ELSNET members to participate in discussion and planning.

Another important conclusion of the panel is that networks need a more principled way of setting priorities, making decisions and measuring results. It recommends that networks draw up a 'road-map', i.e. a document reflecting anticipated progress, world-wide, of the relevant technology and its industrial take-up, as well as the network's place within it. The road-map would define intermediate milestones for the network, thus allowing for a clearer direction and intermediate evaluation of progress. In a community like ours, the design of such a road-map is obviously far from trivial: the integration of language and speech may lead to a variety of results, both in research and in applications, and all of these may require different approaches to technological problems. But I feel confident that with the intellectual power of some 120 organizations in the field behind us, we should be able to accomplish it!

| | |
|--|----|
| Report | 2 |
| An era of medium-scale science? | |
| Announcement | 3 |
| New ELSNET web address | |
| Report | 4 |
| The future of education | |
| Announcement | 5 |
| First ELSNET bullet course | |
| Report | 6 |
| A personal view on East-West Collaboration | |
| Call | 7 |
| for RTD proposals | |
| Conference report | 8 |
| SPECOM'96 | |
| Announcement | 9 |
| ELSNET's 5th European Summer School | |
| New Sites | 10 |
| Announcement | 11 |
| ACL/EACL joint conference | |
| Future events | 12 |

Dec 1996

elsnet
.....

Speech and Language in an Era of Medium-scale Science

Niels Ole Bernsen, Roskilde University

In celebration of the transition from ELSNET-1 to ELSNET-2, I would like to make some remarks on the perennial theme of "Speech and Language: where are we now, what should be done next?". Some of my remarks concern goals that have not (yet) been reached, others developments which ELSNET might want to take into account in charting its future course. I will also make some suggestions for action by ELSNET itself or by others.

ELSNET's long-term Research Goal

The recently drawn-up ELSNET-2 Network Programme describes ELSNET's long-term research goal as follows:

The long-term technological goal which unites the participants in ELSNET is to build integrated multilingual speech and NL systems with unrestricted coverage of both spoken and written language. (ELSNET-2 Network Programme, June 1996)

It is clear that this goal has not been reached yet. This does not mean, however, that no important progress towards it was made during the lifespan of ELSNET-1. Perhaps the single most important achievement in that period was the appearance of continuous-speech, large-vocabulary, speaker-independent speech recognition products. Arguably, the best current example of such a product in Europe is the Philips train time table information system, which allows a two-way dialogue between railway customers and the computer, and which is now being operated by Swiss Rail. The fact that systems such as this one have become a reality by no means implies that all problems in the area of speech recognition have been solved; but it does put an end to the nagging question of when we will finally see advanced interactive speech systems on the market: they are here.

ELSNET's mission, the integration of speech and language, is beginning to come to fruition; European producers of speech technologies are hiring staff as never before.

Medium-scale Science

In 1992, in the early stages of ELSNET-1, a research strategy document drawn up by the ELSNET Research Task group observed that

By focusing on a small number of exemplary tasks and providing shared development data and assessment methodologies for these tasks, a community of effort can be brought about from the bottom up, which will achieve the goals of coherence, cooperation and sensitivity to user needs. (Report by the ELSNET Research Coordination Task Group, August 1992)

The document concludes that research on speech and natural language systems should move away from the isolated three-year-project paradigm towards larger, coherently organised project structures. With the exception of landmark projects such as *VerbMobil*, such a move has not taken place in Europe yet, either nationally or at the European level. At the same time it has become even more clear, however, why such a change is important and necessary: the field of integrated research on speech and language systems has become one of medium-scale science.

The central driving force behind research is the development of research prototypes which are advanced but still realistic from an end-user's point of view. Obviously, this is also desirable from an industrial perspective, because it is hard to think of a better meeting place for industry and academia. This means that in the near future, the problems posed by the development of these systems will increasingly determine the research agendas in the fields of speech and natural language as a whole.

But building such systems is very demanding on resources. The scientific literature typically offers little in terms of guidance, whether in the form of practical system development experience or applicable theory. At the same time, simulation-before-implementation is a necessity, to avoid the risk of

implementing a worthless system concept. Because of this, methods and tools often have to be developed from scratch. Large collections of data on user-system behaviour have to be created, annotated and analysed. Sometimes even the concepts behind the phenomena which need to be annotated, such as speech acts or types of non-cooperative system dialogue behaviour, have to be developed. Methods and tools for advanced data analysis, and system evaluation concepts and methods, are often non-existent or inadequate.

Obviously, it would be impossible for a single three-year project, involving some more senior researchers and, say, four postdocs and/or Ph.D. students, to handle issues of this complexity appropriately, and to develop what they need as they go along. Such projects therefore focus on a subset of issues, losing important progress potential as a result.

Since the onset of ELSNET-1 — and partly at the instigation of ELSNET-1 — relevant European efforts have emerged on evaluation (e.g. *EAGLES*) and resources (*ELRA*, *SPEECHDAT*, *BABEL* and others). These are to be welcomed as part of a much-needed response to the problems pointed out in the ELSNET Research Strategy Paper. But there is no indication that these and other European initiatives are aimed at supporting medium-scale European research efforts, as opposed to the smaller-scale projects we have known for more than a decade.

The Need for Generalization

Another point on which one might like to see innovation, and which deserves strong emphasis in the Fifth Framework Programme, is the development of new theory, concepts, software tools and best-practice methodologies related to the development of advanced interactive speech and language systems. If efficient research progress is to be made, building a new system that goes beyond the state of the art in system building amounts to only half the exercise. The other half lies in generalising the insights gained, into

ELSNews is published at the Centre for Cognitive Science, University of Edinburgh using Aldus Pagemaker™. It is printed on recycled paper by Lutton Press, Ltd.

Editorial team: Mimo Caenepeel (copy editor), Yvonne van Holsteijn (commissioning editor) and Steven Krauwer (managing editor).

Lay-out and production: Mimo Caenepeel

ISSN 1350-990X

© ELSNET 1996.

FOR INFORMATION

Contributions to ELSNews, and address corrections, should be sent to:

elsnews@let.ruu.nl

Fax: +44 131 650 6626

Material for the next issue is due:

15 January 1997

theory, concepts, methods and tools that others may use to produce similar or better systems more efficiently. So far, the 'See my beautiful new system' paradigm appears to have been sufficient for most researchers and funding agencies, no doubt encouraged by the strong emphasis on industrial relevance in Framework Four. But this may not be the best way of exploiting Europe's research potential.

ELSNET and the Web

A further important development which has taken place since the 1992 Strategy Paper is the emergence of the World Wide Web. It is likely that in future many advanced systems integrating speech and language will operate on the Web, and that these will combine speech and language with other modalities for the representation and exchange of information between system and user. Multimodality has arrived and is here to stay.

ELSNET-2 should exploit these new opportunities with determination. There are currently at least two initiatives to that effect being discussed which appear promising. One is the development of an ELSNET 'virtual facility' on the Web, which will provide access to all information, resources, tools, papers and dissertations, training material, software etc. that are publicly available at the ELSNET node sites. Another is the creation of a Web-based electronic infrastructure which could serve as a 'showcase' for speech and natural language technologies; here, ELSNET could take the initiative in developing research projects and bringing together existing systems that will help achieve this goal.

An ELSNET virtual facility would also serve another set of important objectives shared by ELSNET-2 and the Copernicus project ELSNET-Goes-East. Both ELSNET-2 and ELSNET-Goes-East aim to integrate research and industrial activity in Central and Eastern European countries with what happens in Western Europe. A virtual facility and technology showcase mechanisms could play an important role in achieving this goal.

Multimodality and Research Integration

Multimodality is here to stay, and in future many speech and natural language

systems will therefore form components of multimodal systems. This leads to another suggestion, namely that ELSNET might profit from strengthening its links with other relevant European Networks of Excellence for mutual benefit and joint impact. Intelligent multimodal systems, which sometimes include cooperative human-human aspects, are being developed by a multitude of different research communities. These communities include parts of traditional artificial intelligence; intelligent multimedia systems; the speech processing community; the computational linguistics community involved in building systems for text translation, information retrieval, spell and grammar checking etc.; computer vision; computer graphics; acoustics research; computer supported cooperative work, enhanced and virtual reality; and ergonomics. It is likely that these communities, driven by the technological potential for building truly natural interactive computer systems, will come much more closely together in future. And if this is indeed a probable scenario, then ELSNET-2 should consider acting on it.

In the same context, it is worth pointing out that one of ELSNET's goals remains that of integrating, to a meaningful extent, the speech and language communities. I don't think we have ever consulted social scientists on the likelihood of this happening, or on a plausible time scale for it. In any case, it has not happened yet, and

the process is clearly a slow one:

(...) we are still far removed from a situation where it is natural for NL and speech people to work closely together, or even to be aware of each other's research agendas and achievements. (ELSNET-2 Network Programme, June 1996)

The scenario sketched above does not invalidate ELSNET's goal of integrating speech and natural language. But it does draw attention to the fact that the speech and natural language communities are but two communities among many others which share the more general goal of developing natural interactive systems. This could have far-reaching implications for ELSNET-2. We are only just beginning to grasp these implications, and they are likely to influence more and more the course of ELSNET's research strategies during ELSNET-2.

FOR INFORMATION

Niels Ole Bernsen is co-convenor of the ELSNET Research Task Group. He can be contacted at Centre for Cognitive Science Building 3.1.5 Roskilde University, Denmark
Tel: +45 46 75 55 81/ X 2619
Fax: +45 46 75 45 02
Email: nob@cog.ruc.dk

www.elsnet.org comes on-line

During recent months, the ELSNET Information Dissemination TG has been reviewing the network's web site. Although ELSNET can boast of having established a site early on in the mad rush to the WWW (some of our files go back to January 1994!), our site has only undergone small incremental changes since then. As a result, the page structure and the underlying directory structure are both in need of radical reconstruction, the end of which is almost in sight. While the refurbished site won't be cutting-edge 'cool', we hope it will be easy to navigate and quick to load. We have also introduced more data abstraction through the use of (precompiled) Perl scripts, which should make the site easier for us to maintain and update.

In conjunction with this overhaul, we have registered [elsnet.org](http://www.elsnet.org) as a domain with InternIC (before someone else snapped it up). As a result, the URL for the new site will be easier to remember: <http://www.elsnet.org/>.

Another related ELSNET initiative is GLOSSA, a web-based directory of resources (software, data, courses, and so on) in the area of NL and speech. The implementation of the database and interface has been carried out, and we are now adding information. GLOSSA will be available via a link from the new web pages, and has been set up so that organisations can update their own (password-protected) information. Details are available on: <http://www.elsnet.org/glossa/>.

Dec 1996

elsnet
.....

The Landscape of Future Education

Gerrit Bloothoof, Utrecht University

Whither education, and what is ELSNET's (future) role in it? Gerrit Bloothoof, Convenor of the Training and Mobility Taskgroup, weighs up the issues.

One of the world's oldest Institutes of Phonetics was founded in Hamburg at the beginning of this century. Its establishment was made possible through the support of politicians at the time, who were keen to advance the experimental study of language and speech. It is worth considering, however, that their motivations were largely economic: Hamburg was a commercial centre, and the senators felt that the promotion and understanding of foreign languages and cultures would help maximize profits. The pursuit of science and economic considerations were closely intertwined.

In the current climate of budget cuts, economic arguments are surfacing likewise, and are affecting the debate, and decisions, on university education. People (and policymakers in particular) are asking whether students are being educated in a way that equips them to meet the needs of society and industry. In their choice of subject area, students tend to opt for academic disciplines that offer promising employment prospects; and governments, on their part, are becoming increasingly reluctant to fund studies which attract few students. These issues are carrying more and more weight these days, while considerations relating to academic freedom and cultural value appear to be less influential than they used to.

Against this background, a debate about future education is essential for the Language and Speech Sciences. Writing a scenario on how education should develop would be naive — such a directive approach is deemed to fail. But encouraging initiatives that inspire, and that bring to education the kind of dynamics which keep it vital and alert, is productive. And the more successful we are at anticipating new developments, the better our chances for survival. To achieve this, communication about teaching practices, books and papers is essential, as is a continuous discussion with industry, society and academic colleagues to define the professional profile of our students. In future, the Internet will no doubt play an increasingly important role in this respect.

In terms of resources, most research funding in the area of Language and Speech these days is allocated to application-oriented technology projects. Similarly, employers are predominantly looking for people with strong engineering skills — witness the requirements specified in the steadily growing number of vacancies in the Language and Speech industry on *elsnet-list*. But where does this leave university programmes in, say, Arts Faculties, and how should the latter respond to this trend?

Education in the non-applied sciences is certainly not superfluous. A recent questionnaire by Roger Moore revealed that scientists and developers consider basic research in Speech to be of prime importance for technological developments; I

assume the same is true for Language. It is not easy, however, to equip students with an Arts background with engineering skills; and examples of engineering curricula with a strong basic research component in phonetics and linguistics are few and far between. In many countries, the Language and Speech Communication Sciences constitute only part of a broader curriculum, and extensions are hard to implement. At the same time, the implementation of courses of studies where Language and Speech combine to form a single full curriculum is also proving to be difficult. Nijmegen University is one of the pioneers in this respect, and a case in point: it faces difficulties attracting students. This is due to a lack of awareness of the subject area among prospective students: the study of Language and Speech is simply not promoted in secondary education. For this to change, Language and Speech technology will have to develop into a booming market with a high impact on society. But can current departments and research centres survive that long?

Within the ELSNET community, an interest in education-related issues is very much alive. I witnessed this this Spring when I put together a bid for a thematic network under the EU SOCRATES programme, in which reflection on education was a central concern. The bid was successful, and although the proposed network focuses on speech communication, curricular links with the Language Sciences will be on the agenda.

ELSNET has always put great emphasis on education and training, and supports a variety of initiatives in this area. The annual Summer Schools, which are among the major achievements of the network, are probably best-known in this respect. ELSNET has recently applied for funding under the TMR programme to stimulate participation of students from less-favoured regions and less favoured language communities, and of young researchers employed by SME's. The Summer Schools have in turn led to other initiatives. For instance, the first ELSNET book, *Corpus-Based Methods in Language and Speech Processing*, based on the 1994 Summer School, is due to appear soon; and plans are afoot for a second book on Dialogue Systems. In addition to this, ELSNET recently coordinated an application for the development of a European Masters Degree in Language and Speech. If the EU SOCRATES programme funds this proposal, work on it will start in September 1997.

Other ELSNET initiatives in the area of education include *WEBSLS*, the electronic student journal supported jointly by ELSNET, ESCA and EAEL; and the organisation, by ELSNET, of bullet-courses, short intensive courses on topics of vital industrial interest. The first of these, on Dialogue Systems, will be held in Utrecht in April 1997 (for more information see the announcement on the opposite page).

Academics and industrials share an interest in university education: both thrive on people with an open mind and an ability for team work. I very much hope that the new ELSNET



industrial panel will help to establish and consolidate long-term contacts between industry and universities, enabling productive collaboration on curricular development, traineeships, courses and practicals, and so on.

The scope of ELSNET's educational activities has been widening from the very beginning, and that is precisely what is needed: the creation of a landscape of opportunities for high-quality study, training and jobs for our students.

FOR INFORMATION

Gerrit Bloothoof can be contacted at:
 Research Institute for Language and Speech (OTS)
 Trans 10
 3512 JK Utrecht
 The Netherlands
 Phone: +31 30 253 6042
 Fax: +31 30 2536000
 Email: Gerrit.Bloothoof@let.ruu.nl

FIRST ANNOUNCEMENT

ELSNET Bullet Course on Designing and Testing Spoken Dialogue Systems

April 16-18, 1997

Utrecht, The Netherlands

Course
 announcement

For spoken-dialogue systems to be effective in real-world use, dialogues need to be expertly designed and carefully tested. ELSNET's first bullet course will equip participants with skills in this area. It will provide them with basic knowledge of interaction design, which will enable them to develop new applications; and practical knowledge emerging from field experience, which will enable them to create dialogues that work.

Course content:

- Principles of interaction and context in designing dialogues.
- Fundamentals of the dialogue design process, including the application of dialogue theory to specific domains.
- Dialogue management techniques.
- Design and acceptance testing.
- Real-world experience with dialogue systems.

Key points will be illustrated through the building of a prototype dialogue system; this way participants will gain concrete experience, getting expert feedback in the process.

Lecturers

Harald Aust, Philips Research (Germany)
Norman Fraser, Vocalis (UK)
David G. Novick, EURISCO (France)

Who should attend?

The course is intended for potential users of state-of-the-art technology in Spoken Dialogue Systems. These include technology providers as well as companies who want to make their services more efficient or cost-effective (e.g. call centres). The course is geared particularly towards personnel at (sub) management level for whom it is useful to get acquainted with the possibilities (and limitations) of spoken dialogue technologies.

Course Format

This is an international three-day course, consisting of seven 3-hour sessions, starting at lunch time on day one. The course language will be English. Participation will be limited to 25 people.

Course site

The course will be held at the conference centre Oud London near Utrecht, the Netherlands.

Fee

The fee for the course, which includes accommodation, breakfast, lunch and dinner, and documentation is HFL 2950,- (1375 ECU). ELSNET members are offered a 10% reduction. The topic for this course was selected on the basis of a questionnaire distributed to ELSNET members earlier this year. Those who replied to this questionnaire will get a 15% reduction.

Registration deadline

March 3, 1997

FOR INFORMATION

The ELSNET bullet course on Spoken Dialogue Systems is organized in collaboration with OTS (Utrecht University) and DUS (De Utrechtse School). For more information, and to register, please contact:
 ELSNET
 Trans 10
 3512 JK, Utrecht
 The Netherlands
 Tel: +31 30 253 6039
 Fax: +31 30 253 6000
 Email: elsnet@let.ruu.nl

Dec 1996

elsnet

Bridges to the Future

A personal view on East-West collaboration

Erik-Jan van der Linden, ILLC, University of Amsterdam

Erik-Jan van der Linden, coordinator of Elsnet-Goes-East, has been involved for some time in establishing and consolidating contacts with colleagues in Central and Eastern Europe. In the next issue of ELSNews, which will be entirely devoted to East-West collaboration, he will describe new developments in Elsnet-Goes-East, and highlight the problems and opportunities the project has given rise to. Here, he gives a personal perspective on East-West collaboration, and the motivations behind it.

Why be involved in efforts to enhance cooperation between Central/Eastern (CE) Europe and Western Europe? In my own case, concern with such issues arose at least in part from the kind of environment I work in. The ILLC performs a lot of 'community service'; and since it is a little bigger than its sister institutes both within and outside The Netherlands, it tends to take the initiative in, or responsibility for, establishing networks, national graduate schools, associations etc — the type of activities that in EC-lingo would be referred to as 'concerted action'.

Many of my colleagues here in Amsterdam have been involved in collaborative endeavours with colleagues in Central and

conscientious objector) recently went so far as to visit the NATO headquarters in Brussels in order to arrange funding for the 1996 European Summer School in Logic, Language and Information in Prague (no, this is *not* the ELSNET school); I would recommend the experience, by the way. It all goes to show how you can become infected by your environment in these matters...

A second motivation for being involved in East-West collaboration lies in the fact that 'Going East' seems to have a spirit of adventure associated with it. I am aware of this when I hear my colleagues talk about "The first time we met a number of Russians in Sofia in 1988..." and see their eyes beginning to widen slightly. I detect a similar response in myself, and believe that our esteemed network coordinator in Utrecht will testify to the same. The 'adventure' lies partly in coming into close contact with countries and cultures very different from your own; but it also has to do with interacting with people carrying out research like your own in different, and often difficult, environments.

Working in a place where colleagues talk to you daily about these matters, involve you in CE projects, and make you talk to their visitors from Central and Eastern Europe, obviously has its



Eastern Europe since the '80s. When asked, they will give you two reasons for their involvement. The first one is a commitment to their discipline at large. Many groups in Central and Eastern Europe carry out high-quality work, and contact with them is felt to be vital for research in certain areas to flourish. During the crisis in the early '90s, people here tried hard to keep in touch with their colleagues in Central and Eastern Europe, not only out of concern for the scientific survival of their friends, but also, by implication, for that of their discipline.

I believe that this sense of responsibility for their field as a whole distinguishes Western-European researchers from many of their colleagues elsewhere in the world. Often, there is little self-interest in the trouble they take to secure funding for a project or initiative, and their efforts may result in little gain for themselves or their immediate environment. I have seen colleagues spend months talking to research council officials, the European Commission or INTAS-officials, to end up with a mere handful of ECUS — most of it for colleagues in Eastern Europe. And, speaking more personally, I (a former

effects; and over the years I have acquired some experience in this area. In 1994, Elsnet-Goes-East was set up, as a more structured project extending ELSNET-related activities in Central and Eastern Europe; and since then I have been able to put the experience gained in my working environment to a longer-term, and larger-scale, use.

Elsnet-Goes-East has just received a good review, and plans for the future are in the making; in other words, the project is at a stage where it's natural to step back and look at what you've been doing.

Did I infect enough people? I think so. Issues relating to Central and Eastern Europe are now on the official agenda of ELSNET, and get discussed at each Board meeting. As I write, our network coordinator is in St. Petersburg, after having visited several other CE countries as well: I'm happy to have played a part in infecting him. I also hope to have infected my colleagues in Central and Eastern Europe, and made them keen to collaborate with their counterparts in the West. In the past, the latter group has sometimes had an unfortunate tendency to impose their own research strategies upon their CE-colleagues; but this seems to have improved lately.

e

Did I do everything correctly? I don't think so. As in every venture aimed at improving international cooperation, there are lots of things to be learnt. The following conversation, which took place in a car somewhere South of Moscow in November 1995, gives a small sample of this:

Erik-Jan: "Would it be an idea to make an inventory of all linguistic PhD-programmes in Russia, then?"

(short silence in the car)

Igor Boguslavsky: "Erik-Jan, you should remember that Russia is BIG.... Every Russian town with a university has such a programme. There must be hundreds of them. Would it be OK if I limit myself to the programmes in computational linguistics?"

Erik-Jan (nods): "Ehhh ... sure... eh...."

So much for a personal view. In the next issue of *ELSNets*, which will be devoted entirely to East-West collaboration, we will tell you all about ELSNET's Correct Party Line in this area; and we hope to be able to present you with a new slogan. We had the 1994/1995 'ELSNET goes East', and the 1996/1997 'ELSNET goes East, and stays there'. What's next?



Kremlin, Moscow

FOR INFORMATION

Erik-Jan van der Linden can be contacted at:

ILCC

University of Amsterdam

Plantage Muidergracht 24

1018 TV Amsterdam

The Netherlands

Tel: + 31 20 5256090

Fax: +31 20 5255206

Email: erikjan@fwi.uva.nl

The ILCC's web-site address is

<http://www.fwi.uva.nl/research/ilcc>

Preliminary Announcement

Fourth LE call for proposals (RTD projects)

Language Engineering (LE) is the third largest sector of the Telematics Applications Programme 1994-1998. LE-sponsored activities aim to increase the possibilities for human and business communication both in and between European languages, by exploiting new spoken and written language-processing methods. Research and Technology Development (RTD) focuses on pilot projects that integrate language technologies into telematics systems and services.

The Call for proposals was published in the EC Official Journal on 17 December 1996. That means that the proposers' Information Package is now available, along with the application forms, the revised work programme and other related documents.

All Call-related LE documents will be published on the I*M-EUROPE Web server (URL: <http://www2.echo.lu/langeng/en/lehome.html>). The closing date for proposals is 15 April 1997.

FOR INFORMATION

For all general inquiries regarding the Telematics Applications Programme, please contact:

DG XIII-C-1

Help Desk and Proposals Office

Rue de la Loi 200,

B-1049 Brussels

Fax: +32 2 295 2354

Phone: +32 2 296 8962

For all specific inquiries regarding the Language Engineering Sector, please contact:

DG XIII-E-5 LE

Office Batiment Euroforum (0-177)

Rue Alcide de Gasperi

L-2920 Luxembourg

Fax: +352 4301 34999

Phone: +352 4301 32886

Announcement:

LE call

Dec 1996

elsnet
.....

SPECOM'96 in St Petersburg: the end of Russian Capitalism?

Steven Krauwer, *Utrecht University*

I recently attended the International Workshop on Speech and Computer (SPECOM'96), organized in St Petersburg by the State Pedagogical University and the St Petersburg Institute for Informatics and Automation. It was an event which can be viewed as a stepping stone towards more and better collaboration between Language and Speech researchers all over Europe. In what follows I will present a number of statistics relating to various aspects of the workshop; I would argue that these illustrate some more general characteristics of current East/West collaboration that are worth reflecting on.

First, the content of the workshop. The Proceedings consist of 41 papers, covering a range of topics organized in terms of the various sessions at the workshop. Some 15 of the papers address issues on the borderline between Speech and Language (or solely Language); this suggests that the organizers share ELSNET's views on the integration of Language and Speech.

Second, the workshop's profile in terms of the nationality of organizers and speakers. The international nature of the event is reflected in the composition of the Scientific Committee (which had six members from Russia and Belarus, and six from the the European Union), and in the national affiliations of the (first) authors of the papers (11 from Russia, five from Rumania, five from outside Europe, and 20 from the EU). These numbers show that the balance in terms of overall East-West participation was about right. Note, however, that the number of Rumanian papers was relatively high, while there were no papers at all from other Central and Eastern European countries. I have no explanation for this; it may be due to existing collaboration patterns (or rather the lack of). Due to the international character of the workshop the working language was English; but this clearly

constituted a stumbling block for some of the Russian participants.

As far as collaborative ventures are concerned, there were eight papers by authors from two or more different institutions. Three of these were the product of East/East collaboration (one within Russia, two within Rumania), three of West/West collaboration, and two of East/West collaboration (Rumania with Spain and France, respectively). There was no collaborative work between Russia and EU countries, which is disappointing; hopefully future events of the SPECOM'96 kind will lead to more joint work in this respect.

In terms of bibliographical references, there were some significant differences between Eastern and Western papers. The following table illustrates this; it gives the total number of references for all the papers in the proceedings, as well as the number of references to the author's (or authors') own work, and to (other) Eastern or Western sources:

| Papers | Number | | | Percentage | | |
|-------------|--------|------|------|------------|------|------|
| | All | East | West | All | East | West |
| Ref to Own | 82 | 47 | 55 | 26 | 43 | 19 |
| Ref to East | 18 | 18 | 0 | 51 | 16 | 0 |
| Ref to West | 273 | 45 | 228 | 69 | 41 | 81 |
| Total Refs | 393 | 10 | 283 | 100 | 100 | 100 |

Bibliographical references in SPECOM'96 papers

First of all, the average number of references per paper is seven for the Eastern-European papers, and 11 for the Western ones. This difference is easily explained: it is extremely difficult for many of our colleagues from Eastern Europe to get hold of books and journals published in the West. This also explains the proportion of references to Western literature in particular: viz. 81% in papers from the West, as opposed to 41% in the Eastern-European ones.

Two other differences, however, are more striking. The first one is the complete

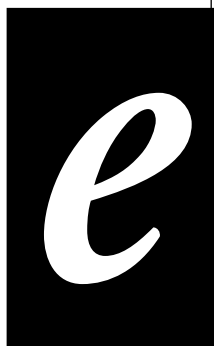
absence of references to Eastern work in the Western papers. I can think of at least three possible explanations for this, viz (i) lack of quality; (ii) lack of relevance; and (iii) limited accessibility of the Eastern European literature. Explanation (i) I refuse to believe in; (ii) does not strike me as improbable, as work done in Eastern Europe cannot always be easily connected with what counts as mainstream research in Western Europe and the US; and (iii) seems likely, because a relatively high number of Eastern-European research publications are written in Russian, and limited in their distribution

Another notable difference highlighted by the table above concerns in the high number of self-references by Eastern authors. While Western authors refer to their own work in 19% of the cases, for Eastern-European contributions this is 43%. I think that this figure reflects that Eastern researchers have a tendency towards individualism, and prefer to create their own scientific paradigm. This seems

to be in line with what I said above about the potential lack of relevance for Western researchers of at least part of the research carried out in the East. Let me emphasize that I am not expressing, or implying, any value judgements in this respect; but it is something to keep in mind for future collaborations.

The observations presented here are based solely on the SPECOM'96 workshop, so we should be careful not to overgeneralize; they do, however, fit in with my impressions of similar events I have attended over the last two years.

What conclusions can we draw from all this? For me, the most important one is that there seems to be a clear determination amongst our Eastern colleagues to establish and maintain working relationships with researchers in NLP and Speech all over Europe. But ensuring better access to our research



output in Eastern Europe — e.g. by sending paper or electronic copies of books, papers or journals printed in the West — seems to be crucial. I would also recommend that colleagues from Eastern Europe make an effort to improve their active and passive knowledge of English, so that the exchange of ideas and results in both directions can take place more smoothly. And finally, I would argue that events like *specom'96* (small, mixed workshops with ample opportunity for interaction) constitute a very effective way of enhancing scientific integration between East and West, and that they should continue to be organised, not only in Russia but in all countries in Central and Eastern Europe.

Finally, with reference to the title of this article, how could *SPECOM'96* herald the end of Russian capitalism? One last feature of the event worth drawing attention to is that it was not organized or hosted by an institution in Russia's capital. Capitals tend to be a natural first port of call in establishing contacts with the research community in a country; but this should not lead to a situation where they automatically end up dominating contacts. *SPECOM'96* has clearly demonstrated that a successful workshop need not take place in a capital (although I am aware of the fact that St Petersburg did serve as the capital of Russia for a long time). I hope that we will see many more non-capitalist events in the near future.

FOR INFORMATION

SPECOM'96 was jointly organized by the St Petersburg Institute for Informatics and Automation, of the Russian Academy of Science, and the State Pedagogical University of Russia. Copies of the proceedings can be obtained from the organizers: contact kosarev@specom.spb.ru.

Steven Krauwer can be contacted at ELSNET

OTS, Utrecht University, Trans 10
3512 JK Utrecht, The Netherlands

Tel: +31 30 253 6039

Fax: +31 30 253 6000

Email: elsnet@let.ruu.nl

WWW: <http://www.elsnet.org>

ELSNET'S FIFTH EUROPEAN SUMMER SCHOOL

Lexicon Development for Language and Speech Processing

July 14-25, 1997

Centre for Computational Linguistics, KU Leuven, Belgium

Announcement:
ELSNET Summer
School

Language and Speech Processing applications with a reasonable coverage need large lexicons. The construction and maintenance of lexical resources, however, is costly and complex. Because of this, decisions about notation, format, structure, content and acquisition of lexical knowledge require careful consideration, with due attention to specific requirements of applications, and reusability factors.

Students and staff interested in the development of systems for Speech and Language Communication need to be familiar with the latest developments in this area. This year's Summer School will provide a full picture of the current state of the art, during two weeks of plenaries, parallel courses and practical workshops in Leuven, an attractive historical town in Belgium.

Topics to be covered include:

- computational lexicography for Speech and Language
- supporting & automating lexicography
- use of lexica in text-to-speech systems
- speech databases
- knowledge bases for NLP
- syntactic aspects of lexical structure
- lexical semantics
- use of lexica in Automatic Speech Recognition

The Summer School is open to advanced undergraduate students, postgraduate students, postdocs, and staff members

from academic and industrial sites. The number of participants will be limited to 90. Participants are expected to have a basic knowledge of Language and/or Speech Processing, but no specific prior experience in the area of lexicons will be assumed.

The programme committee consists of **John Nerbonne** (Groningen University) and **Gerrit Bloothoof** (Utrecht University). Local organisation is coordinated by **Frank van Eynde** (KU Leuven).

The Summer School Fee is 125 ECU for students, 250 ECU for staff, and 500 ECU for researchers from industry. Participants from ELSNET sites will get a 20% reduction. As the number of places is limited, early booking is strongly recommended. Student accommodation will be available at appr. 225 ECU for 13 nights (breakfast not included).

Registration deadline: May 1, 1997

Payment deadline: June 1, 1997

FOR INFORMATION

For more information please contact the ELSNET Summer School Secretariat

KU Leuven, Centre for Computational Linguistics

Maria-Theresiastraat 21

B-3000 Leuven

Tel: +32-16-32.50.88

Fax: +32-16-32.50.98

Email: ess97@ccl.kuleuven.ac.be

<http://www.ccl.kuleuven.ac.be/ess97/ess97.html>

Dec 1996

elsnet
•••••

Nokia Research Centre

Peter Boda, Nokia Research Centre

Nokia Research Center is the corporate research and development unit of the Nokia Group. Its technology areas include telecommunications, audio-visual signal processing, software and electronics. The main focus is on wireless communications and multimedia.

The Research Center interacts closely with the R&D units of all Nokia business groups. It covers the full range of activities from exploring new technologies and product/system concepts to their exploitation in actual product development. The Center also introduces new methodologies in Nokia's business processes, and takes part in various international R&D projects.

The Speech and Audio System Laboratory has six groups:

- Speech Coding (standardisation of state-of-the-art speech coding)
- Audio Coding (standardisation of state-of-the-art audio coding)
- Speech and Audio Enhancement (development of methods for wireless telecommunications systems)
- Acoustics (enhancement of speech and audio quality in various Nokia products)

- Low-bit Rate Multimedia (development of solutions for multimedia applications in wireless environments)
- Speech Recognition (enabling flexible interactions in speech-driven applications by means of voice-controlled user interfaces)

The main focus of the Speech Recognition group is naturally on Voice Dialling, but research is not limited to this: it extends to other domains which may be relevant to enhancing user-friendly interactions in speech-driven applications.

FOR INFORMATION

Peter Boda, Senior Research Engineer at Nokia, can be contacted at:

Speech and Audio Systems Laboratory

Nokia Research Centre

Sinitaival 6

FIN - 33720 Tampere

Finland

Tel: +358 3 272 5467

Fax: +358 3 272 5897

Email: peter.boda@research.nokia.com

Natural Language Processing Group

University of Ulster

Norman Creaney, University of Ulster

The NLP Research Group at the University of Ulster is involved in a range of research projects, including:

- Repairing Errors in Computer Assisted Language Learning (RECALL): aims to develop more adequate and more user-oriented software for CALL, incorporating diagnosis and user modelling. (CEC, Language Engineering)
- Language Acquisition in Conditions of Variable Input: aims to establish the effect of variability in the input data of language learning. (ESRC)
- Telematics Enhanced Language Learning and Tutoring Systems (TELOS): aims to implement and demonstrate a telematic service for language learning and tutoring. (CEC, Telematics)

- Spoken Dialogue Systems: The group is involved in several projects involving dialogue control for spoken dialogue systems. Most work in this field at present is based on application-specific solutions. Our aim is to develop a generic application development framework that will identify information elements and processing algorithms common to a wide range of application areas.
- User Modelling: Building on previous work in this area, this group is examining the application of user modelling techniques to language teaching (see RECALL project above) and to the selection of output media in a multimedia information support system.

FOR INFORMATION

Norman Creaney, coordinator of the NLP Research group, can be contacted at:

Faculty of Informatics

University of Ulster at Coleraine

N Ireland

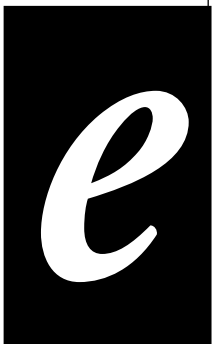
BT52 1SA

Tel: +44 1265 324502

Fax: +44 1265 324916

Email: n.creaney@ulst.ac.uk

http://www.infcl.ulst.ac.uk/~norman/nlp_grp.html



CONFERENCE ANNOUNCEMENT

ACL-97/EACL-97 Joint Conference

Universidad Nacional de Educacion a Distancia (UNED)

Madrid, Spain

July 7-10, 1997

In 1997, the ACL will hold its annual conference for the first time in Europe, in conjunction with the bi-annual EACL conference. The meeting will be co-chaired by **Phil Cohen** (Oregon Graduate Institute, USA) and **Wolfgang Wahlster** (DFKI, Germany). Local arrangements are coordinated by **Felisa Verdejo** (UNED, Spain).

Theoretical, practical, and empirical papers are invited on all aspects of computational linguistics, including (but not restricted to) the following:

- 1) Morphology, Lexicon, and Finite State Technologies
- 2) Grammar and Formalisms for Parsing and Tactical Generation
- 3) Semantics, Pragmatics and Discourse
- 4) Uses of Language Processing
- 5) Statistical Language Processing

There will be a special poster session on innovative work-in-progress by students. Suggestions for workshops, tutorials, demonstrations and additional events are welcome. ELSNET will play an active role in this conference, and we look forward to your participation in making this a stimulating event.

Deadlines

Electronic submissions : January 8

Hard copy submissions : January 10

Notification of acceptance: March 20

Camera-ready copies of papers: May 1

FOR INFORMATION

Detailed information on submission requirements, workshops, tutorials and student sessions, and local organization can be found at the ACL web site address: <http://www.cs.columbia.edu/~acl>

Future Events

Jan 8-10, 1997: *2nd International Workshop on Computational Semantics*. Tilburg, The Netherlands. For information contact Computational.Semantics@kub.nl. URL: <http://tkiwww.kub.nl:2080/tki/Docs/IWCS/iwcsappl.html>

Jan 10-12, 1997: *Socrates/Erasmus Thematic Network "Speech Communication Sciences"*. Kick-off meeting of working groups. For information, contact bloothoof@let.ruu.nl

Jan 13-24, 1997: *LOT Winter School*. Nijmegen, The Netherlands. For information contact lot@let.ruu.nl. URL: <http://www.wots.let.ruu.nl/LOT/ws97.html>

Jan 22-24, 1997: *Fifth International Symposium on Social Communication*. Santiago de Cuba, Cuba. For information, contact lingapli@ceniai.cu

Mar 12-14, 1997: Conference on Audio- & Video-based Biometric Person Authentication. Crans-Montana, Switzerland. For information, contact Pigeon@tele.ucl.ac.be. URL: <http://www.tele.ucl.ac.be/IMAGES/AVBPA>

March 12-14, 1997: *Fuzzy-Neuro Systems 97: Computational Intelligence — Fourth International Workshop*. University of Paderborn, Germany. For information, contact fns97@ibm18.uni-paderborn.de (FNS 97) URL: <http://www.uni-paderborn.de/~fns97/>

April 4-5, 1997: *ANLP-9*. Washington D.C., USA. For information, contact grishman@cs.nyu.edu. URL: <http://cs.nyu.edu/cs/projects/proteus/anlp97>

April 16-18, 1997: *ELSNET Bullet Course on Designing and Testing Spoken Dialogue Systems*. Utrecht, The Netherlands. For information, contact elsnet@let.ruu.nl

Apr 17-18, 1997: *Robust Speech Recognition for Unknown Communication Channels*. Pont-a-Mousson, France. For information, contact p.green@dcs.shef.ac.uk. URL: <http://www.dcs.shef.ac.uk/~pdg/esca-ws/esca-ws.html>

Apr 17-20, 1997: *TELRI: Second European Seminar*. Kaunas, Lithuania. For information, contact telri@ids-mannheim.de. URL: <http://www.ids-mannheim.de/telri/telri.html>

Apr 20-23, 1997: *Sixth Annual Symposium on Document Analysis and Information Retrieval*. Alexis Park Resort, Las Vegas, Nevada. For information contact frank@little-charlie.ISRI.UNLV.EDU

Future events

Dec 1996


 The logo for ELSNET, featuring the word "elsnet" in a stylized, lowercase font with a series of dots underneath it.

ELSNET Participants Academic Sites

NL Utrecht University (coordinator)
 A OFAI/Univ. Vienna/Vienna Univ. of Technology
 B University of Antwerp
 B University of Leuven
 BU Bulgarian Acad. of Sciences, Sofia
 BY Bellarussian Academy of Sciences, Minsk
 CH IDSIA, Lugano
 CH ISSCO, Geneva
 CZ Charles University, Prague
 D Univ. des Saarlandes/DFKI, Saarbrücken
 D Univ. Hamburg
 D Univ. Kiel
 D Univ. of Stuttgart
 D Ruhr-Univ. Bochum
 D Univ. Erlangen
 DK Ctr for Sprogteknologie, Copenhagen
 DK Ctr for PersonKommunikation (CPK), Aalborg
 DK Ctr for Cognitive Science, Roskilde Univ.
 E Universidad de Granada
 E Univ. Politecnica de Catalonia/Univ. Autonoma de Barcelona
 E Univ. Politecnica de Madrid
 E Univ. Politecnica de Valencia
 F LIMSI-CNRS, Orsay
 F IRIT, Toulouse
 F Inst. de la Comm. Parlée, Grenoble
 F IRISA, Rennes
 F Laboratoire Parole et Langage-CNRS, Aix-en-Provence
 F CRIN, Nancy
 GR ILSP/NCSR "Demokritos", Athens
 GR Wire Communications Lab., Patras
 H Hungarian Acad. of Sciences, Budapest
 H Technical University, Budapest
 I Ist. di Linguistica Computazionale, Pisa
 I IRST, Trento

I Fondazione Ugo Bordoni, Rome
 IRL University College Dublin
 IT Institute of Mathematics and Informatics, Vilnius
 N University of Trondheim
 NL Stichting Spraaktechnologie, Utrecht
 NL Inst. for Perception Research, Eindhoven
 NL Leyden Univ.
 NL Catholic Univ. of Nijmegen
 NL TNO Human Factors Reseach Institute
 NL Univ. of Amsterdam
 NL Univ. of Tilburg
 P INESC/ILTEC/Univ. Nova de Lisboa
 PL Polish Academy of Sciences, Warsaw
 RO Research Inst. for Informatics, Bucharest
 RU Russian Academy of Sciences, Moscow
 S KTH, Stockholm
 S Univ. of Linköping
 UK Defence Research Agency, Malvern
 UK UMIST, Univ. of Manchester
 UK Univ. of Cambridge
 UK Univ. College London/School of Oriental and African Studies (SOAS)
 UK University of Edinburgh
 UK Univ. of Essex
 UK Univ. of Dundee
 UK Univ. of Leeds
 UK Univ. of Sheffield
 UK Univ. of Sussex
 UK Univ. of Ulster
 UK Univ. of York

D Langenscheidt
 D Novotech GmbH
 D pc-plus Computing
 D Philips Research Laboratories
 D Siemens AG
 DK Jydsk Telefon
 E Telefonica I&D
 F ACSYS
 F Aerospaciale
 F GSI-ERLI
 F LINGA s.a.r.l.
 F MemoData
 F Rank Xerox Research Center
 F Systran SA
 F TGID
 F VECSYS Speech Processing
 FI Nokia Research Center
 GR Knowledge A.E.
 H Morphologic
 I CSELT
 I Database Informatica
 I Sogei (IRI-FINSEIL Group)
 I Syntax Sistemi Software
 I TecnoPolis CSATA Novus Ortus
 NL KPN Research Laboratories
 NL Polydoc N.V.
 P Uninova CRIA
 RU Analit, Ltd.
 S Telia Promotor (Call Centre Division)
 UK ALPNET UK, Ltd
 UK BICC plc
 UK British Telecommunications
 UK Cambridge Algorithmica Ltd.
 UK Canon Research Centre Europe Ltd.
 UK Enigma Ltd.
 UK Hewlett-Packard Labs
 UK Logica Cambridge Ltd.
 UK Sharp Laboratories
 UK SRI International
 UK Vocalis Ltd.

Industrial Sites

B Lernout & Hauspie Speech Products
 D aspect GmbH
 D CAP debis
 D Daimler-Benz AG
 D Electronic Publishing Partners GmbH
 D Grundig Professional Electronics GmbH
 D IBM Deutschland

ELSNET

Secretariat

Steven Krauer
Coordinator

Yvonne van Holsteijn
Assistant Coordinator
Utrecht University

Task Group

Convenors

Training & Mobility
Gerrit Bloothoof, Utrecht University

Info Dissemination
Ewan Klein
Edinburgh University

Industrial Links
Anne de Roeck
University of Essex

Linguistic & Speech Resources
Antonio Zampolli
Istituto di Linguistica Computazionale

Research
Niels Ole Bernsen
Roskilde University
and
Joseph Mariani
LIMSI-CNRS

What is ELSNET?

ELSNET, the European Network in Language and Speech, was established in 1991, with funding from ESPRIT Basic Research. There were 25 founding members of the network. Currently, there are more than 60 universities and research institutes, and more than 45 companies participating.

The long-term technological goal which unites the members of ELSNET is to build integrated multilingual NL and speech systems with unrestricted coverage of both spoken and written language. Building multilingual NL and speech systems requires a massive joint effort by two pairs of communities: on the one hand, the natural language and speech communities, and on the other, academia and industry. Both pairs of communities are traditionally separated by wide gaps. It is ELSNET's objective to provide a platform which bridges both gaps, and to ensure that all parties are provided with optimal conditions for fruitful collaboration. To achieve this, ELSNET has established an infrastructure for sharing knowledge, resources, problems, and solutions by offering (information) services and facilities, and by organising events which serve academia and industry in both the language and speech communities. In this respect, it is important to note that a network like ELSNET can only function well if all members of the network are prepared to give and to receive.

Electronic Mailing List

elsnet-list is ELSNET's electronic mailing list. Email sent to elsnet-list@let.ruu.nl is received by all Managing, Associate and Industrial node coordinators of the Network, as well as other

person who are not necessarily members of ELSNET, but who have an interest in ELSNET's activities. This mailing list may be used to announce activities, post job openings, or discuss issues which are relevant to persons in the European natural language and speech communities. To request additions/deletions/changes of address in the mailing list, send mail to luz@cogsci.ed.ac.uk.

ELSNET's WWW Document

Detailed information about ELSNET and its activities and publications is available through the Web at the following URL:

<http://www.elsnet.org>

Comments and suggestions for new WWW pages are very welcome. In particular, each ELSNET site coordinator is encouraged to send details of his or her site's home page so that a hyperlink might be set up to it from the ELSNET home page.

FOR INFORMATION

ELSNET

OTS, Utrecht University, Trans 10
3512 JK Utrecht, The Netherlands

Tel: +31 30 253 6039

Fax: +31 30 253 6000

Email: elsnet@let.ruu.nl

WWW: <http://www.elsnet.org>