

# Teaching and learning computational linguistics in an international setting

**Koenraad de Smedt**

[www.uib.no/acohum](http://www.uib.no/acohum)

[acohum@uib.no](mailto:acohum@uib.no)

## Summary

Computational Linguistics has long been a forerunner in the use of humanities computing technology. However, there are many organisational problems to be addressed to maintain and improve the quality of teaching and learning of Computational Linguistics. Advanced Computing in the Humanities (ACO\*HUM) is an international network which investigates the use of new technologies in Humanities teaching and learning. It promotes international co-operation a.o. for teaching and learning Computational Linguistics.

## Aims

In so far as Computational Linguistics belongs to the human sciences or humanities, it has long had a privileged position. It has long been a forerunner in the application of formal methods and of computational techniques from Artificial Intelligence. It has during the past twenty or so years established itself as an advanced field of research and study, and it has transformed both computer science and linguistics. However, its special position among humanities disciplines is rapidly vanishing. The sciences of language are today not at all unique in their adoption of advanced computing. History, for example, is rapidly absorbing computing as an important research methodology, for example for the intelligent searching of archives. History of art is rapidly adopting advanced visual processing techniques, the building of pictorial databases, etc. Even literature, which has long been a place of traditional scholarship based on books, is increasingly using computer-aided research techniques, for example in formal stylistics, and is beginning to take hypertext serious as basis for a new rhetoric. And with Sophie's world on CD-ROM, even philosophers are trading their pens for shiny disks. Briefly, the whole humanities are today faced with a challenge to innovate their learning and teaching.

Back to Computational Linguistics, what are the challenges Computational Linguistics is facing? The field of natural language processing has expanded enormously in the past decades and is continuing to expand with respect to theory, methods, and applications. This has allowed many different schools and many niches of special research. While this variety is useful and needs to be maintained, it is not so easy in everyday life to offer quality teaching in Computational Linguistics covering an up to date part of the Computational Linguistics spectrum. For one thing, the Computational Linguistics staff at most universities is rather small. For another, the development of good teaching materials is way behind the research developments. This may affect the quality of teaching in such ways that students are not sufficiently familiar with what research requires, students are not prepared for computational linguistics professions, and students do not learn some things which are considered basic at other universities, which results in the hindrances for their mobility. All this argues for increased co-operation, also on

international levels, to develop and offer better teaching and training in Computational Linguistics.

These challenges are food for thought for the SOCRATES / ERASMUS thematic network on Advanced Computing in the Humanities (ACO\*HUM). What do we envisage? Part of the efforts of ACO\*HUM is directed towards curricula innovation. We aim at curricula containing a range of basic modules with a wide international agreement, while allowing for local specialisation's. Here we build on seminal work done by the former ERASMUS network on natural language processing. Another part of our efforts is directed to agreeing on equivalencies of degrees, and possibly the creation of a new international masters degree in natural language processing. We also aim at the development of new teaching materials for Computational Linguistics, including web-based teaching. Web-based courses offer not only the opportunity to learn across national boundaries, but also to teach across national boundaries. Several individuals in the Computational Linguistics community are experimenting, but lack up till now a forum for the exchange of experiences. ACO\*HUM is actively stimulating the development and testing of more such courses, in co-operation with ELSNET. We intend to use the network as a forum for exchanging experiences on a trans-national basis. In this context we also stimulate also the transfer of research results to teaching materials, so that today's students learn to work with systems and data which are the state of the art. Finally, we promote Computational Linguistics as a professional profile, and we want to co-operate with professional organisations such as EACL and with research institutions and companies to study the correspondence between what students learn and what the real world expects from them as professionals.

Not only Computational Linguistics, but also mainstream linguistics is affected by the developments. Mainstream scholars of language are increasingly using tools which are produced by Computational Linguistics research. One needs to distinguish here between the linguist, who is a user of linguistic tools, and the computational linguist, who is a developer of linguistic tools. Even without understanding how a parser works, a parser can be a practical tool for a linguist who wants to play with grammars. This argues for the inclusion of tools courses in linguistics curricula. In fact, at the University of Bergen, all first year linguistics students use such tools as the LFG workbench and Tarski's world. A co-operation between linguists and computational linguists is needed in the updating of linguistics curricula, and also here a co-operation is useful. In fact, ACO\*HUM does not aim at an isolation of those using advanced computing in the humanities and those who don't. Rather, we want to bring advanced computing to as large a part of humanities students as possible.

### **Organisation of the network**

Advanced Computing in the Humanities (ACO\*HUM) is an international network project which investigates the use of new technologies in Humanities teaching and learning. The project started in September 1996 and will continue for 3 years. It is a network project, which means that it is based on the mutual exchange of information and on voluntary agreements. The network is no regulating agency, but offers a forum for exchange among the partners. This exchange takes the form of working groups which regularly meet, workshops and conferences. In September, a large conference will be organised in Bergen; this will be announced on the web and announcements will be sent to Nodali. Through voluntary co-operation in the network, the

project aims at facilitating mobility including virtual mobility and common educational resources. Eventually we hope to find a common ground for international degrees in Computational Linguistics, international pooling of educational materials and computational training materials for Computational Linguistics, and better contact between educational institutions and future employers.

The network has an office based in Bergen. They can be reached at [www.uib.no/acoHum](http://www.uib.no/acoHum). Currently, the network has six working groups which are represented schematically in Figure 1.

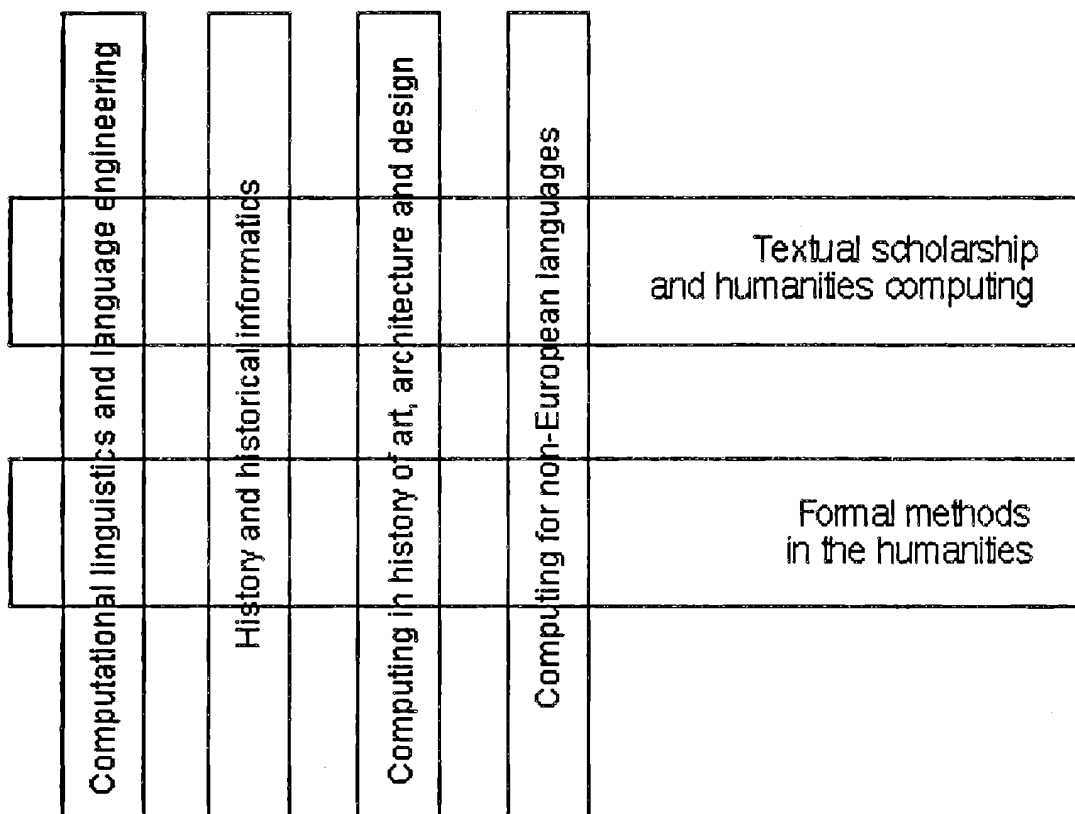


Figure 1. ACO\*HUM working groups

As shown in the figure, there is an interplay of vertical themes and horizontal themes. Vertical themes correspond largely to traditional disciplines within the humanities, while horizontal themes are humanities wide areas of interest.

The working group on Computational Linguistics and language engineering is composed of Bill Black and Marie Hayet (Manchester), Walter Daelemans (Antwerpen and Tilburg), Laurence Danlos (Paris), Koenraad de Smedt (Bergen), Gert Durieux (Antwerpen), Joakim Nivre (Göteborg), Hans Uszkoreit and Brigitte Krenn (Saarbrücken), Paul Mc Kevitt (Ålborg), Julia Lavid and Felisa Verdejo (Madrid), Torbjørn Nordgård (Trondheim) and Andy Way (Dublin). These members have been active in setting up discussions and actions related to the issues mentioned earlier, and will by the end of the project produce a report and recommendations.

## **Preliminary conclusions**

The whole project has been active for one and a half year, which is relatively short for a large network with over a hundred partners. The conclusions can therefore only be preliminary. Our partners have up to now indicated that within several humanities disciplines, clear trends are noticeable based on real student needs. Among these trends we mention the following:

1. There is increasing demand for collaborative practice. This is perceived to be a radical shift away from the cult of the individual towards the development of a new perception of the individual practitioner as a member of a team, or part of a network of relationships, both within and across disciplines. Information and communication technology is perceived as instrumental for new collaborative models. Students, also in Computational Linguistics, have to learn how to communicate and work together with others.
2. Graduates in Computational Linguistics are badly wanted on the job market in many regions of Europe, a.o. in Scandinavia. However, students need to be equipped with up to date knowledge which is needed in the real world, not just classical theories and methods. This means they need access to good course modules and up to date learning materials.
3. The use of information and communication technology should be applied to teaching and learning situations in order to improve the efficiency and quality of academic education. Bringing the computer into teaching in a reasoned way should liberate students from time and space limitations on learning, enable life-long and distance learning as well as augmenting traditional degree schemes. Concerted actions are needed to create technical and organisational conditions for such developments on an international basis.