

**Integration of the project “Information Technology
in Computer-Assisted Translation of Croatian
and in e-Language Learning” into curriculum**

Sanja Seljan

Department of Information Sciences

Phone: +385-1-6002350, Fax: +385-1-6002431, E-mail: sanja.seljan@ffzg.hr

Nataša Pavlović

Department of English Language and Literature

Phone: +385-1-6120055, Fax : +385-1-6120234, E-mail: natholdet@gmail.com

Faculty of Humanities and Social Sciences - University of Zagreb
Ivana Lučića 3, HR-10000 Zagreb, Croatia

Summary

This paper aims to describe the integration of the project for an under-resourced language, Croatian, into curriculum at two departments at the Faculty of Humanities and Social Sciences, University of Zagreb, especially for students of Information Sciences and English. In particular, we would like to share our plans to incorporate the three-year project entitled “Information Technology in Computer-Assisted Translation of Croatian and in e-Language Learning” into the translator training curriculum. The project has started as a response to changes at the national level and the need for integration into European surroundings. The basic aim of the project is to investigate and to develop, as much as possible, resources and tools for CAT and for eLL for Croatian, according to the experience of new EU member states and in accordance with EU standards.

Keywords: technology, language, machine, computer, translation, e-learning, curriculum, project, EU standards, Croatian.

1. Project

This project, named CroCAsTeLL (**C**roatian **C**omputer-**A**ssisted **T**ranslation and **e**-**L**anguage **L**earning), involves an interdisciplinary team of experts from the fields of information sciences and language and translation studies (English, French, Croatian), with members from both the academic setting

and from the industry. Foreign partners from universities in Slovenia, Austria and Italy, as well as from a government institution in Slovenia are also participating.

The research will involve insights from various models that already exist for other languages, starting with the alignment process and building of translation memory (TM), terminology bases and morphologically sensitive modules, proceeding with statistical machine translation, example-based machine translation, simulations studies and models related to the introduction of the new language into existing models (Seljan, 2006).

In addition to the (morphologically sensitive and insensitive) statistical analyses, which have not been done on Croatian texts yet, methods will be examined for word, phrase and sentence alignment necessary for the building of translation memory systems based on parallel corpora. Lexical and syntactic relations between Croatian and English relevant for the building of machine translation systems (EC Systran) will be examined. The research will be made in highly specialized areas and in accordance with European standards.

Another part of the project will concentrate on the use of the existing models in the area of computer linguistics and language and translation didactics. In cooperation with a small, privately-owned company, this project aims to be integrated into the curriculum at graduate and postgraduate levels through several different courses. First, we will say a few words about the educational context, and describe what has been done so far.

2. Translator training at the Department of English

At the present time, in Croatia there are no pre-graduate programs in translation, and most future translators are educated at departments of modern languages (about 70% of all those engaged in translation and/or interpreting, according to a survey reported in Pavlović forthcoming). We will here describe the situation at the Department of English at the University of Zagreb, which is the oldest in Croatia. Additionally, we will say something about what is being done at

the Department of Computer Science at the same university, and about our plans to further the cooperation between these two departments in relation to translator training.

Faculty of Humanities and Social Sciences at the University of Zagreb consists out of 23 departments, offering close to 80 pre-graduate programs to 6,100 students, as well as a number of graduate and postgraduate programs. Until now, students who wished to become full-time or part-time translators took a four-year B.A. degree program in two languages, after which they could take a one-year vocational M.A. in translation. At the moment a reform is under way as part of the Bologna process. Under the new system, which is now entering its third year, students take a general three-year course in two languages leading to a first (B.A.) degree, after which several second-cycle tracks will be on offer, among them a two-year M.A. program in translation. Students can also take an additional one-year course in interpreting, which was launched in 2006.

Students of English are expected to have a relatively high level of competence in that language even at the time of the enrolment (candidates are selected on the basis of an entrance exam; the ratio of candidates to available places is generally 7:1). By the time our students get to their fourth year of studies, they have generally been learning English for at least twelve years. Their L2 competence at the end of their fourth year is expected to have reached the equivalent of level C (“proficient user”) of the Council of Europe’s (2007) *Common European Framework of Reference, CEFR*.

2.1. Changes in the translation course

The practical translation classes (“exercises”), especially translation from Croatian into English, were originally devised with a view to improving the students’ knowledge of English as a second language. However, the explosion in demand for translation on the Croatian market starting from the early 1990s has meant that a large number of our students end up working as full-time or part-time

translators (and/or interpreters). To meet this demand, the practical translation class has undergone a transformation: from a language-learning venue, in which translation was a means to an end, it has gradually changed to a translation-learning venue, insofar as this has been possible under the time constraints. This means that we work with authentic texts, including occasional authentic translation projects, and considerations such as translator resources, target readership, task specifications, fees, translator ethics, translation norms, etc. are discussed regularly. Professional translators and major employers are invited several times a year to provide students with additional real-world perspectives. Students have the opportunity to take a week of translator traineeship at the Ministry of Foreign Affairs and European Integrations, a translation agency, or Croatian Television (HRT).

The main objective of the practical translation courses is to help students acquire the foundations of translation competence – the skills needed to “produce an acceptable target text in one language on the basis of a text written in another” (Király 2000: 13) – as well as translator competence – “being able to use tools and information to create communicatively successful texts that are accepted as good translations within the community concerned” (2000: 13-14).

In this context, the integration of language and translation technology into the training program is obviously a burning issue. We now describe some aspects of this integration that we have already incorporated in our classes.

2.2. E-learning system

In 2004, the Faculty of Humanities and Social Sciences introduced an e-learning platform (Stančić, 2007), Omega (<http://omega.ffzg.hr/>), based on Moodle (www.moodle.org). This has enabled us to complement our class work with online activities and resources. As far as our pre-graduate translation workshops are concerned, this means that the students download their assignments from the website, and also upload their translations to a designated place. The

teacher then looks at their work and provides feedback, which is also made available on the website. The Forum feature enables the students to exchange links and tips, as well as to discuss the translations they are working on. Students can also compile their own glossaries of specialized terminology using the Glossary feature.

Our main problem has been the large number of students and the small number of contact hours for translation workshops, as our students have a wide array of other courses they are taking. The e-learning platform has made it possible to expand our class work beyond the limitations of the time we have available for the face-to-face sessions. It has also enabled the teachers to gain a better insight into the difficulties particular students are facing, and to offer guidance and advice. This has relieved some pressure from the face-to-face sessions, which are now freed up for collaborative student work on projects, and similar activities.

One of the problems we are still facing is the fact that in the pre-graduate programs teachers are not paid for online work. It is our hope that in the near future this situation will change for the better.

2.3. Collaborative learning

Another thing we have introduced over the past few years is a trend to incorporate collaborative learning (cf. Kiraly 2000) into our translation workshops. This has shifted the focus from a teacher-oriented to student-oriented classrooms, and encouraged the students to become more independent in their approach to both learning and translation. As we suggest elsewhere (Pavlović 2007), collaborative learning seems to be a very useful method for L2 translation classes, which are important in settings in which a “language of limited diffusion” is used.

So far, we have tried one authentic translation project, in which the students were asked to translate from English into Croatian the users’ guide for Moodle. The “client” and terminology specialist was the administrator of our e-

learning website. In the future, we would like to focus even more on authentic projects for real clients, including those outside the university. Our main problem, in addition to the large number of students and small number of contact hours is the lack of equipped translation labs.

3. Curriculum changes at the Department of Information Sciences

The curriculum changes implemented at the Department of Information Sciences (Seljan, Pavuna, 2006; Stančić, 2007) are the result of educational and organisational changes in line with the policy statements in Croatia and the Bologna process, but also respecting the EU standards regarding use of ICT in higher education system, the need for a better alliance with industry and professional practice, of cooperation with foreign universities, national and international projects, internal and external evaluation processes, mobility programs and the need for approaching knowledge society.

3.1. Study programs at the Department of Information Sciences

At the Department of Information Sciences, students can choose either a program consisting of two majors (in combination with a modern language) or as a single major, both of which are based on an interdisciplinary approach achieved through modules and elective courses. Apart from the general knowledge from the field of information sciences such as the basics of ICT, networks, programming techniques, mathematics, databases and information retrieval, logic and mathematics, specialist knowledge is also offered. Students can choose among the different courses related to natural language processing, formal languages, language databases, computational linguistics, lexicography, knowledge management, digitisation process and e-learning.

Courses that are offered to all students of Information Sciences but also to other students of the Faculty, in which the results of project work are to be implemented are the following: *Machine Translation* in the 3rd year of the pre-

graduate program, and *Computer-Assisted Language Learning, Authoring Tools in e-Language Learning* and *Computational Grammar Models* at the graduate (second-cycle) level. At the graduate level the two courses that are directly related to the project are *Computer-Assisted Translation* and *ICT in Language Learning*. Cooperation with the Department of French has already started through joint research of multimedia use in language learning, and also with the Department of English through their *Translation Workshop*.

3.2. Courses related to the project

- The course *Machine Translation (MT)* is adapted to numerous European and American courses, and to EU recommendations (DGT, 2007). It includes a theoretical perspective and practical work related to use of all types of available translation tools (online dictionaries, thesauri, terminology bases, translation memories, machine translation software). The theoretical part deals with the development, applications, limitations, needs, levels of automation, strategies of translation, linguistic problems, role of controlled languages and evaluation. In practical work, students are asked to do the exercises that more or less follow the lectures, such as the use of electronic dictionaries, thesauri and terminology bases (available on the Internet), comparison of different translation software mostly available on Internet (Systran, Globalink, ProMT, etc.), practical work regarding the use of translation memories (Trados, WordFast, Atril), analysis of mistakes and software evaluation, aiming to cover the use of tools recommended by the EU (DGT 2007), pre- and post-editing and to present the translation workflow. Students are asked to work in pairs and to present in class some research of interest related to current state of machine translation. All presentations, related papers and activities are given in the traditional way and by e-learning solution. Seminars, usually presented as .ppt, are uploaded individually by students.

At the beginning of the course students are given a questionnaire regarding their previous knowledge of MT/ CAT and at the end of the course they

are asked for their opinions and impressions. While their previous knowledge is rather scant (mostly because of the lack of Croatian language resources), the final results have shown satisfaction especially because of the practical work.

Experience in the use and creation of smaller translation memories acquired in the project will be further transferred into the course, as well as alignment problems, building of terminology bases, ambiguity problems, and part of language analysis using different software. This part of the practical work will be done in part by the students of English and of Information Sciences in a computer lab, depending on the type of software available.

- The course *Computer-Assisted Language Learning* includes a combination of theory, the evaluation of existing programs, practical use of ICT in language learning and relevant methodological aspects. The theoretical part encompasses the historical development, overview of language technologies, role of multimedia, skill development using ICT and evaluation criteria. Practical work follows the lectures, using all types of electronic resources independently of the language (from dictionaries to specialized educational software, fun software, authoring tools, text processors, exams, plans, etc.) serving as a complementary tool to the traditional way of teaching. Students are also asked to present the research done in pairs related to possible application of a certain topic, specific type of software evaluation and possible use or creation of a web page related to a certain topic with integrated cultural elements.

Experience acquired through the project regarding the use of multimedia as a complementary tool, online assessment, role of feedback, specific terminology from a certain area or specific syntactic structures, etc. will be presented and used in class.

- The course *Authoring Tools in e-Language Learning* aims to show possible applications of authoring tools in the language-learning class, along with evaluation, methodological aspects, and ICT and language skills independently of

the language in question. Students will create all types of tests and quizzes including multimedia elements in order to practice language skills and using feedback information. Special attention will be given to standards and technological aspect of authoring tools, using traditional and e-learning systems. Students will work in pairs and present group of exercises related to a certain topic in a suitable context or create a web page using different types of authoring tools. Experience acquired through the project related to terminology could be used in this class.

- In *Computational Grammar Models*, attention will be paid to formal models that could be used to parse certain syntactic (terminological) constructions. After considering the role of formal models in language technologies, attention will be given to models that are implemented in the process of machine translation or language learning, in combination with other approaches. Special attention is given to the Lexical-Functional Grammar (LFG) model where students would be able to develop smaller parsers for the segments of the Croatian language that could resolve certain terminological constructions or formal models analogue to other languages.

4. Conclusion

Results of such an interdisciplinary approach integrating professional practice in the academic context, and answering to educational, professional and organizational changes, is expected to provide students with a broader understanding of the area, creating flexible professionals, equipped for teamwork and problem solving. Working on real problems and with real texts, relating theory to practice, students have so far shown to be very satisfied and willing to work. It is expected that the integration of this project into the curriculum will also meet with a favorable response from the students of both departments.

For under-resourced languages, such as Croatian, it is important to integrate research projects as much as possible into the European context, in

particular regarding the curriculum changes, the building of electronic language and translation resources, project cooperation, adoption of standards and good strategy development.

Acknowledgement

The presented results are the outcome of the research project “Information Technology in Computer-Assisted Translation of Croatian and in e-Language Learning” (130-1300646-0909) undertaken with the support of the Ministry of Science, Education and Sport of the Republic of Croatia.

References:

- Belam, J. Teaching Machine Translation evaluation by assessed project work. Proceedings of the 6th EAMT Workshop, Manchester, 2002.
- Cavalli-Sforza, V.; Brown, R.D.; Carbonell, J.G.; Jansen, P.J.; Kim, J.D. Challenges in Using an Example-Based MT System for A Transnational Digital Government Project. Proceedings of the 9th EAMT Workshop, Malta, 2004.
- Clavier, V.; Poudat, C. Teaching Machine Translation in non Computer Science Subjects: Report of An Educational Experience within University of Orléans. MT Summit VIII Workshop, Santiago de Compostela, 2001.
- Council of Europe. CEFR, Common European Framework of Reference, 2007. (http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf)
- Directorate-General for Translation of the European Commission. Translation Tools and Workflow, 2007. http://ec.europa.eu/dgs/translation/bookshelf/tools_and_workflow_en.pdf
- Forcada, M. L. A 45-hour Computers in Translation course. MT Summit IX Workshop on Teaching Translation Technologies and Tools; 3rd Workshop on Teaching Machine Translation, New Orleans, 2003.
- Forcada, M.L. Learning machine translation strategies using commercial systems: discovering word-reordering rules. MT 2000: Machine Translation and Multilingual Applications in the New Millenium, Exeter, 2000.
- Gaspari, F. Teaching Machine Translation to Trainee Translators: a Survey of Their Knowledge and Opinions. Workshop on Teaching Machine Translation. Santiago de Compostela, 2001.
- Kiraly, D.C. A Social Constructivist Approach to Translator Education. Empowerment from Theory to Practice. Manchester: St. Jerome, 2000.
- Mitamura, T.; Nyberg, E.; Frederking R. Teaching Machine Translation in a Graduate Language Technologies Program. MT Summit IX Workshop on Teaching Translation Technologies and Tools, New Orleans, 2003.
- Moodle. <http://www.moodle.org/>
- Pavlović, N. Directionality in Collaborative Translation Processes. A Study of Novice Translators. Unpublished PhD thesis submitted for eval. at Universitat Rovira i Virgili, Tarragona, Spain, 2007 (Abst. <http://isg.urv.es/publicity/doctorate/research/pavlovic.html>)

- Pavlović, N. Directionality in translation and interpreting practice. Report on a questionnaire survey in Croatia (Forthcoming)
- Pérez-Ortiz, J.A.; Forcada, M.L. Discovering Machine Translation Strategies Beyond Word-for-Word Translation: a Laboratory Assignment. MT Summit VIII, Santiago de Compostela, 2001.
- Seljan, S. Information Technology in Machine Translation and in e-Language Learning of Croatian. Proceedings of the 1st International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006. Badajoz, Spain, vol. II, 2006, 359-363.
- Seljan, S.; Pavuna, D. Why Machine-Assisted Translation (MAT) Tools for Croatian? Proceedings of 28th International Information Technology Interfaces Conference, ITI. Cavtat/ Dubrovnik, 2006, 469-475.
- Stančić, H.; Seljan, S.; Lasić-Lazić, J. Digitisation and Language Technologies in the Learning Process of Information Sciences – Approaching the EU Standards. Proceedings of CIMQUSEF, Morocco, 2007 (in print)