

Teaching Theory and Tools

Translation Technology and Software Localization at the University of the Saarland

by Karl-Heinz Freigang

There is a long tradition in research in the field of language technology at the University of the Saarland in Saarbrücken, Germany. As early as the late sixties, linguistic research projects were launched to analyze German syntax and morphology based on large corpora of German scientific and encyclopedic texts. In the early seventies, research activities began to concentrate on the field of machine translation (MT). The first MT system that was developed at the University of the Saarland was a custom version of the SYSTRAN Russian-English MT system to German as a target language. Soon these activities lead to the development of the first prototype of the SUSY system (Saarbrücker UebersetzungsSYstem) for the language pair Russian-German which was later on expanded to the language pairs French-German and English-German.



Karl-Heinz Freigang

During the last few years of these research activities, ending in 1986, the project team from the Department of Applied Linguistics, Translating and Interpreting also included in its objectives investigations into the working environment of translators. Topics considered included the prospective developments in the field of machine-aided translation, the necessary re-organization of workflow in various translation environments caused by the use of tools, and the necessary re-design of curricula for the training of future translators.

Research project on integrating language technology in a translation curriculum

Based on these activities, a research project "Linguistic data processing as a component in the training of translators and interpreters" was launched in April 1988 and carried through until March 1993. The goal was to develop a model for updating the degree programs for translators and interpreters by incorporating a language technology component into the existing curriculum. Initially, different translation workflows in various translation environments were investigated, in order to better understand the needs and perspectives of translation practice.

The research project produced a curriculum which still forms the basis of the training in the field of language and translation technology at the Department of Applied Linguistics, Translating and Interpreting at the University of the Saarland.

After a period of consolidation of the training program and the establishment of a division for language technology within the Department, the curriculum is now on its way to be officially integrated in the obligatory study schedule and examination regulations of the diploma degree program for translators and interpreters. In contrast to similar programs at other universities, the key elements of this program are designed as an integral part of the regular training program for translators and interpreters, which has to be attended by all students of the degree program.

In contrast to similar programs at other universities, the key elements of this program are designed as an integral part of the regular training program for translators and interpreters, which has to be attended by all students of the degree program.

Language Technology Program

The program includes the following key elements:

- Introduction to computer technology and natural language processing, taking into account the particular requirements and conditions in the field of translating and interpreting, including data communication, Internet use, etc.
- Introduction to computer-aided terminology work including the basic principles of terminology theory and the underlying principles and practical application of terminology management software.
- Introduction to machine-aided translation and machine translation including the basic underlying linguistic and technical principles and concepts as well as practical application of translation tools and machine translation systems.
- Introduction to project management and translation workflow in various translation environments.

These key elements are closely linked to the "traditional" components of the training of translators and interpreters, particularly in the field of LSP (Language for Special Purposes) translation.

Since 1998 the Department of Applied Linguistics, Translating and Interpreting has established its own computer laboratories with 26 PCs (Windows NT 4.0) connected in a local NT network and connected to the university network and the Internet. In these two laboratories (one with 21 PCs, one with 5 PCs) most of the translation tools available on the market are installed, including TRADOS 5, Transit, IBM Translation Manager, Déjà Vu, Trans Suite 2000, T1, Personal Translator, Power Translator, and Systran Enterprise, as well as the localization tools Catalyst, Passolo, and RC WinTrans. All translation software publishers provided their software to the department free of charge or at reduced prices. In addition to these translation tools, the most widely used standard office applications are also available on all workstations, including Microsoft Office, Corel WordPerfect Suite, FrameMaker, Ventura Publisher, and Adobe Acrobat.

While the overall framework of the curriculum has been maintained, the content of the key elements is changing rapidly

from semester to semester. Currently, they are:

1. Introduction to computer technology and natural language processing

This class combines theoretical issues and practical computer exercises. Even today new students need to be introduced to the basics of computer hardware, data representation and data encoding as well as operating systems, since surveys carried out regularly every year among beginners show that there is still a majority rating their own



computer literacy as very low. Based on this basic knowledge, other topics include the basic principles and methods of Natural Language Processing and different areas of language technology (like speech understanding, character recognition, parsing mechanisms), a deeper understanding of concepts of word processing systems and a detailed

The first course in software

localization at the Department was offered in winter term 1993/94.

introduction into Internet technologies (e-mail, search engines, Web design, etc.).

2. Introduction to computer-aided terminology work

This class is also designed as a combination of theoretical lectures and practical com-

puter exercises. It provides an overview of the basic methodology of terminology

All translation software publishers

provided their software to the department free of charge or at reduced prices.

work, the different approaches to computer-aided terminology management, a market survey on terminology software, a detailed discussion of terminological data categories and, of course, practical exercises with terminology management systems. In this course students are also encouraged to create their own terminology databases to be populated and used in the LSP translation classes.

3. Introduction to machine-aided translation and machine translation

This course, which is developed in close cooperation with the department's chair of Machine Translation, also consists of a combination of theoretical lectures and practical exercises. After an overview of various linguistic and computational concepts of machine translation, the major systems available on the market or currently under development are presented, e.g. Systran Enterprise, Logos, Power Translator, Personal Translator, and T1. Besides these machine translation systems, this class also provides an introduction to computer-aided translation tools with an overview on the basic principles underlying translation memory systems.

4. Introduction to translation workflow and project management

Since knowledge of workflow and organization of the translation process in different environments is increasingly in demand by future employers of the students of translation institutes, this course is increasingly important within the curriculum. The course presents an overview of the working organization in different translation environments, from freelance translator to team workers in translation or localization companies. As an example, the German DIN 2345 standard ("Translation Contracts"), which describes the most important aspects of translation projects which have to be negotiated between translator and customer, is discussed in detail. In practical exercises students learn

to create their own database for the management of customer and project data and to calculate translation projects taking into consideration varying constants and variables. Besides using standard database and spreadsheet applications, like Microsoft Access and Excel, special project management software for translation purposes is presented.

This framework of four key elements in the field of computer aids for translators is offered regularly in each study year.

Special Seminars and Optional Classes

More specialized courses are offered to those students who want to gain a deeper insight in translation technology or who want to specialize, for example, in the field of software localization. These additional courses are either offered as seminars or as optional courses where students can obtain a participation certificate.

In each semester a special seminar is offered, either in the field of terminology theory and management or in the field of translation technology. The objective of these seminars is either a detailed investigation into methodological problems of terminology management or a detailed

analysis of new software products in the field of translation. In recent years, seminars were organized analyzing different terminology systems (like Multiterm, TermStar, Termbase included in SDLX, System Quirk) or investigating methods and principles underlying the alignment of bi-lingual texts including an analysis of alignment tools like WinAlign or the

Since knowledge of workflow and organization of the translation process in different environments is increasingly in demand by future employers of the students of translation institutes, this course is increasingly important within the curriculum.

alignment component of Transit. Although these seminars demand a high degree of independent work in getting acquainted with the software to be analyzed, they are very popular and the number of partici-

pants has always to be limited to twenty students in order to guarantee an efficient working environment for the participants.

Classes in Software Localization

Optional courses are offered on a regular basis on "Translation with Translation Memory Tools" and on "Software Localization". The first course in software localization at the Department was offered in winter term 1993/94. The course was organized as a project with weekly meetings of small student working groups with at least one teacher present and "plenary" sessions every three weeks lasting up to three or four hours where all students and all participating teachers met in order to discuss intermediary results of the working groups. The aim of this project was the localization into English and French of a Windows-based terminology management software application developed in German by a former student of the department (TERMISYS for Windows). Ten students took part in the project accompanied by four teachers (two from the language technology division, one from the English division and one from the French language division). Six students were responsible for translation into English, the others for



Karl-Heinz Freigang

translation into French. In both language groups, translation of the online-help could not be finished, but at least the most important components of online-help systems and the relevant file formats could be shown and discussed. Translation (localization) of the software user interface was done in Borland Resource Workshop, the manual (in Ventura Publisher format) and some parts of the online-help were translated in Word (ASCII-text exported from Ventura and RTF files).

The second software localization project was organized in winter term 1999/2000 within a seminar, which meant that not only practical problems and tasks had to be dealt with but also theoretical and methodological issues of software localization, e.g. different text functions and linguistic characteristics of software strings were discussed. In this seminar 20 students, studying English as their first or second language, participated, meeting in plenary sessions every two weeks and in small working groups every week. The 20 participants were organized in four working groups. After four introductory sessions in which each of the four groups had to present a paper summarizing theoretical literature on software localization, practical work on the localization project began in the computer laboratory. The localization tool which was used for translation of the user interface was Corel Catalyst 2.5, which had been purchased at special rates for education purposes. The German distributor of Corel Catalyst, SAM Engineering, Mühlthal, Germany, estab-

lished a connection between the seminar group and Corel Ireland and thus made it possible that Corel provided a localization kit for the seminar, treating the working group just like a localization vendor with the obvious exception of deadlines and

The quite remarkable interest of the students in all classes dealing with translation technology, even if these classes are not obligatory in their curriculum, shows that they are quite aware of the growing importance of these fields for the translation profession.

payment. So the task of the four student working groups was to localize Corel Catalyst using Corel Catalyst. Of course, the aim of the seminar was not really to produce a finished localized version of Corel Catalyst, but rather to get acquainted with the localization tool and with the tasks and problems which arise in localization projects. At the end of the term, parts of the software were localized, the glossaries built during the localization process were exported into a Multiterm database, and were used in the following semester in a

translation class dealing with the translation of the Corel Catalyst online help.

This localization project is currently being completed by one of the participating students who is writing her diploma thesis on this project, discussing on the one hand the theoretical, linguistic and technical challenges in software localization and finishing, on the other hand, the translation of the software and the online help. For her thesis she was also provided with a temporary license of the new version of Catalyst (now ALCHEMY Catalyst) by SAM Engineering and will include an evaluation of this tool in her thesis.

Conclusion

In the current summer term there is again an optional course offered on software localization, this time with 10-15 students working with RC WinTrans, which offers a direct interface to TRADOS. The quite remarkable interest of the students in all classes dealing with translation technology, even if these classes are not obligatory in their curriculum, shows that they are quite aware of the growing importance of these fields for the translation profession.

This also becomes evident when we consider the still rapidly growing interest of professional translators in the continuous training seminars on translation technology which have been offered at the Department of Applied Linguistics, Translating and Interpreting for more than ten years.

Karl-Heinz Freigang obtained his Diploma as a Translator from the University of the Saarland (English, Russian, German) in 1971. He was closely involved in research work in machine translation from 1977 to 1988, and has been head of the Division of Language Technology at the Department for Applied Linguistics, Translating and Interpreting since 1988. Karl-Heinz was Vice-President of the BDÜ (Bundesverband der Dolmetscher und Übersetzer) and since 2000 he is one of the editors of the MDÜ (Mitteilungen für Dolmetscher und Übersetzer). He has also worked for several years as a freelance translator in the field of software localization. Karl-Heinz can be contacted at kh.freigang@rz.uni-sb.de.