Training translators to localize

FRANK AUSTERMÜHL

University of Auckland, New Zealand

Abstract. *The role that localization training plays in translator training programs is scrutinized with particular reference to the School of Applied Linguistics and Cultural Studies at the University of Mainz in Germany. Specific examples are given of the three types of courses that are run under the umbrella of localization, namely translation for localization (emphasis on software and websites), electronic tools for translators, and theoretical issues of localization. A four-level model of translating for localization is introduced encompassing an introduction to localization and computer systems followed by website and software localization and finally the research component. The importance of real-world examples and project work is emphasized throughout, with particular reference to organization and management and familiarity with industry tools and, at the higher level, the value of thorough research.*

Introduction

This position paper will mainly focus on how localization training should be integrated into translator training programs. Most of the paper will be based on my own approach to website and software localization, as implemented within the constraints of a traditional four-year program at the School of Applied Linguistics and Cultural Studies at Germersheim (University of Mainz), Germany.

Before discussing approaches to localization training within and beyond traditional programs of translation and interpretation, I should say that I consider translation to be an integral and central part of localization, but I also consider localization to be more than translation. Localization comprises several tasks that are traditional translation tasks like, yes, translating text itself, terminology mining and management, or revision. However, there are also tasks like software development and engineering, desk-top publishing or the editing of graphics files, that, at most, lie on the peripheries of technical translation. The boundaries between translator tasks and localizer tasks are rather fuzzy, but they nevertheless exist, as the rather clear division of labor between freelance translators and translation or localization project managers shows. The differentiation between translator tasks and localizer tasks also gives rise to upward mobility schemes for freelance translators.

Given the different textual and technical natures of software products and websites and taking into account their different development and publication cycles, software and website localization should not be lumped together as "just the same kind of localization" but should be considered as different phenomena. The development of comparative text typologies, both for hypertexts as well as for software texts, will help to define the differences between these two localization types. They will also shed light on such hybrid products as computer games or DVDs.

Integrating localization into traditional training programs

With regard to ways of integrating localization into translator training programs I will focus on three areas (see Figure 1). The first, which could be called translation for localization, takes place within traditional translation practice classes and focuses on software and website localization. The second type deals with electronic tools for translators. These courses are usually optional.
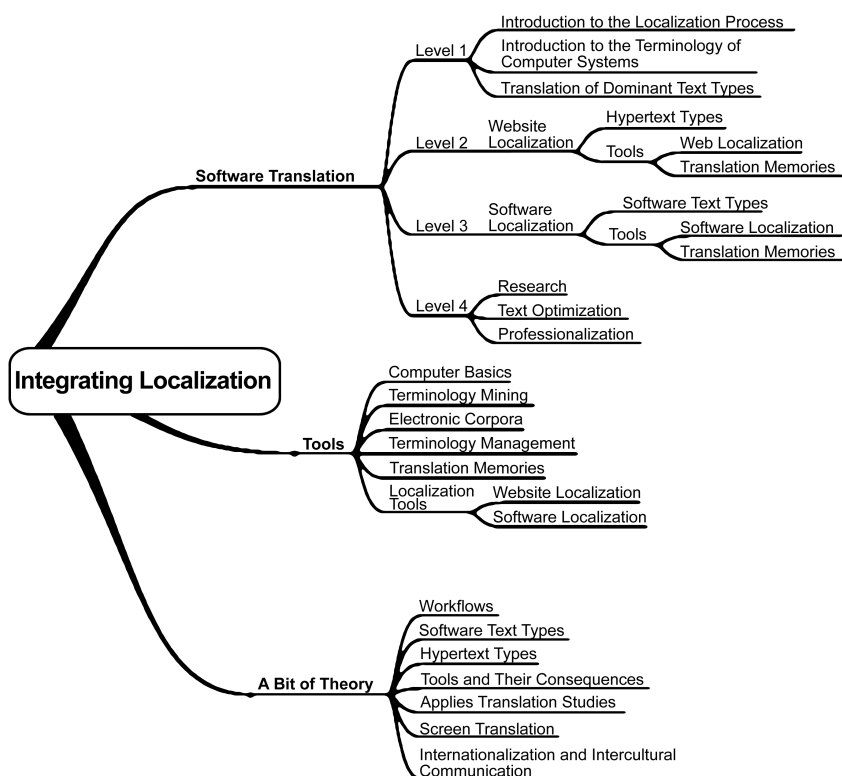

Figure 1: Integrating Localization—General Approach

The third part of the general approach on introducing localization regards seminars dealing with theoretical issues of localization, e.g. workflow analyses, text typologies or translational constraints derived from, for example, the use of content management systems, translation memories, or localization tools. One of the goals of these more theory-oriented courses is to find ways of applying existing translation studies paradigms (for example *Skopos* theory or Holz-Mänttäri's Theory of Translational Action) to localization. Courses in the theory part of the model also deal with issues of internationalization. Given the fact that a good part of the intercultural component of software and website translation is actually located on the level of internationalization, students—as future experts for intercultural communication—learn to apply their bi- or multicultural skills to software programs or multilingual websites.

My thoughts on the question of what elements of localization should be included in a curriculum for translator training and how they should be taught will be guided by the question of feasibility, i.e. what kind of solutions are possible given existing curricular, administrative, and institutional constraints. Accordingly, I will be talking mainly about solutions within traditional translation programs (i.e. the German *Diplom* or the Spanish *Licenciatura*), while solutions outside the traditional scheme of translator training (like specialized postgraduate or MA programs) will only be briefly mentioned. However, since the following proposals are based on a modular approach to localization training, they could easily be used in both of these settings.

Translating for Localization

As stated above, I propose a conceptual differentiation between "translation tasks," i.e. software translation and website translation and "localization tasks" (knowing, of course, that a clear distinction is not possible). The first set of these tasks will most likely be carried out by freelance translators, while the localizer tasks will be carried out by in-house staff, e.g. project managers (a lot of them former freelance translators), software engineers, or quality-control specialists. If we look at localization from the point of view of freelance translators (most of our graduates will use the sector as a first stepping stone into the market) localization could be easily integrated into existing translator-training programs. This could be even more straightforward if we did not start from the notion of localization, which in my mind often places too much emphasis on the *production* process of a multilingual website or software product, but if we instead simply talked about translating software texts and translating hypertexts, something that can be dealt with in rather traditional translation practice classes on technical or scientific texts (only computer-based and -supported, as well as team- and project-oriented).

With regard to my own classes on technical translation, over the course of the past eight years I have developed a four-level model for software and website translation (see Figure 2).
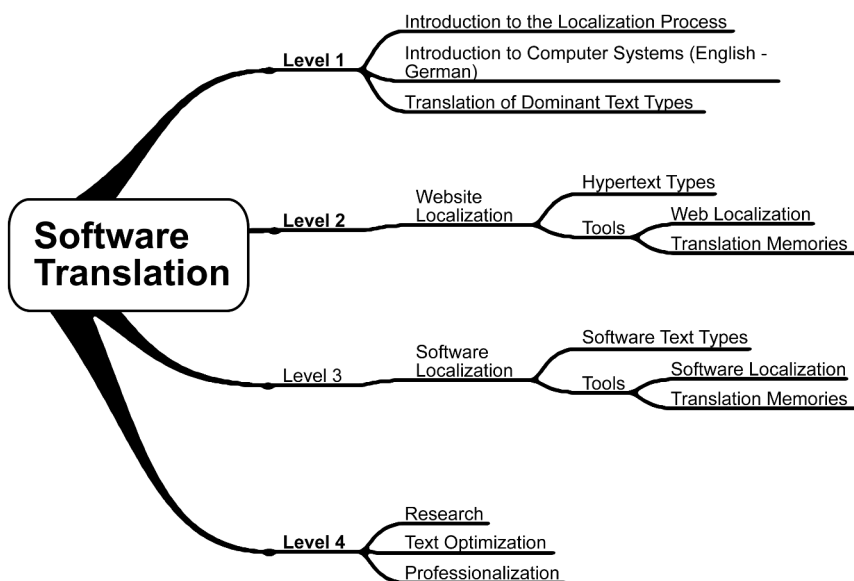

Figure 2: Translating for Localization—Four-Level Model

Each level now comprises about 30 hours. On the first level, students (usually early in their third year) are introduced to localization in general and to the basic text types involved in software localization (on-screen texts, installation guides, Help files). In addition, this course also serves as an introduction to basic hardware and software terminology. This introduction is based on a number of interrelated concept systems (see figure 3 for an example) and fortified through the contents and foci of the texts used throughout the translation practice class (e.g. excerpts from a printer manual or an installation guide for a network adapter).

The courses on the second and third levels (third and fourth years) of the two-year model aim to introduce the participants to "real-life" translation projects. Students not only use typical translator tools such as terminology management systems, translation memories, and localization tools, but also learn how to manage and coordinate small localization projects. On each of the two levels, students carry out a specific localization project, one focusing on software localization, the other on website localization. If possible, these courses are based on real translation tasks, i.e. involving real clients and the subsequent publication of the project results. Where this it not possible, a

real-life project is simulated. The courses include all stages of a localization project from analyzing the source text, calculating the (unfortunately fictitious) budget, organizing and managing the distributed translation of the files, creating and maintaining a project terminology base, building customized corpora, and using CAT tools such as Catalyst or Passolo for software localization or Cats Cradle or Trados Tag Editor for HTML/XML files. Students take on individual roles and become project managers, terminologists, translators or revisers. Software localization projects also comprise the translation of Help files and printed documentation, which can also include the handling of translation-memory systems. For time reasons, however, the use of translation-memory systems usually has to be reduced to a short presentation or has to be left out altogether. Nevertheless, translation memory systems are taught in a different class, and this class is especially geared to third- and fourth-year students.
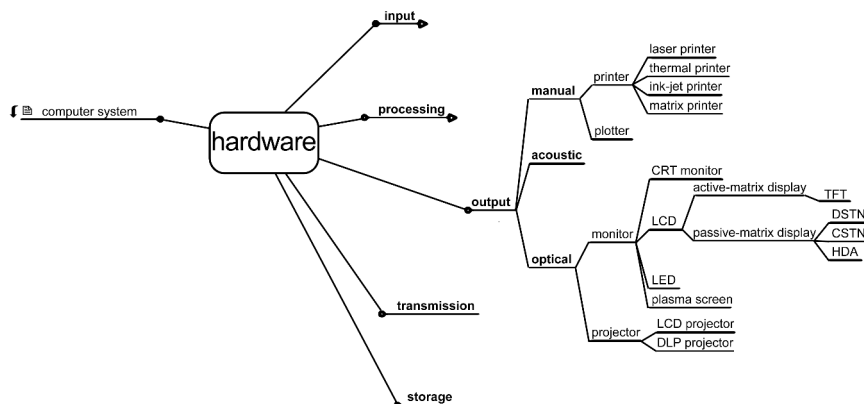

Figure 3: Concept system "Hardware—Output Devices

The fourth level of this translating-for-localization model is directed at exam candidates, usually in their fourth year of study. The final exam in technical translation is a three-hour written translation of a five-hundred-word text. The translation is written by hand and no electronic resources are allowed, but students are allowed to use a print copy of a glossary that they themselves have put together during the semester leading to the exam. An exam course is necessarily influenced by the nature of the exam and student's hopes of and expectations, e.g. solid knowledge of the subject area dealt with and the terminology involved as well as confidence in analyzing and producing culturally-adequate texts. From the student's point of view, this course is very much about "panic control". In addition to this primary goal, the course aims at developing some crucial professional skills. One of these skills is the ability to research a topic thoroughly (I usually pick a

rather new and/or unknown topic) and to build a strong knowledge base that includes the main terminology and phraseology in the field. Starting from this knowledge base, which is supported by a terminology database and a customized bilingual corpus, students are encouraged to self-confidently create independent texts. The source texts usually represent technical marketing material or detailed product descriptions and force the students to dig deep into their encyclopedic and terminological repertoire while at the same time leaving them more room for creativity than, for example, a set of software strings or a Getting Started Guide does. In addition, these types of texts also encourage the students to work on their revision skills. Class discussion is usually based on one sample translation. This presentation is prepared by a group of three students, of which one serves as terminologist, one as translator, and one as reviser.

All four courses mentioned above are obligatory, and are supported by a course website and either a mailing list or a newsgroup.

On Translator Tools and Localizer Tools

The four levels making up the "translating-for-localization model" is supported by a number of additional, non-obligatory courses on electronic tools for translators. The teaching of these courses is divided into several categories, as shown in figure 4.
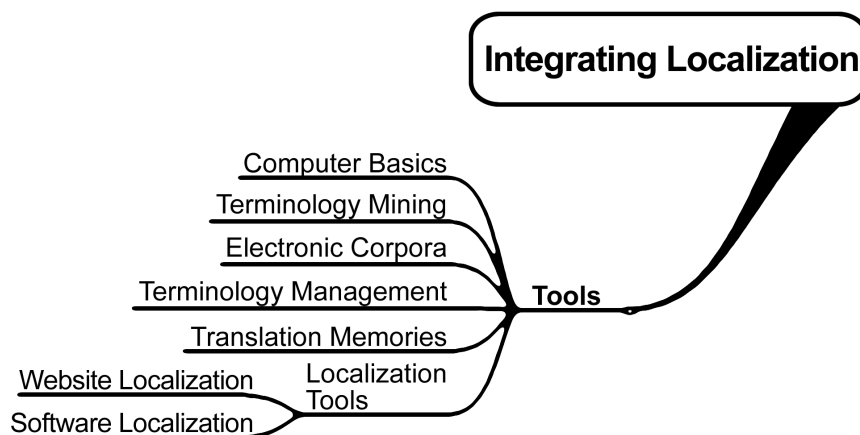

Figure 4: Integrating Localization—Tools

In general, we distinguish between two separate, yet closely intercon-nected sub-processes, each requiring its own set of task-specific tools. The first of these domains comprises the "classical" three-step translation model of source text reception, information transfer, and target text formulation. The computer-based resources used during this core translation process aim

at providing the translator with the linguistic, encyclopedic, and cultural information necessary to successfully perform the task. Since we consider translation to be an utterly knowledge-based activity (see Stolze 1992), these "translator tools" will ideally serve to enhance the translators' hermeneutic abilities, thus allowing them to unfold their full creative potential.

This ideal situation of a translator's freedom, however, is in many cases torpedoed by a second group of electronic tools. These applications, which we will call "localizer tools", aim primarily at streamlining the business process of translation, especially with regard to larger, repetitive translation tasks and projects. Although from the point of view of a human translator it is tempting to characterize these tools (primarily translation memories or localization tools) as merely productivity-enhancing, their impact on the improvement of translation quality, especially with regard to terminological and phraseological consistency, should not be ignored.

Here we focus on the various translator tools, on translation memory and localization solutions, which are used by translators and localizers alike, and on machine translation systems. The other tools listed are primarily used by larger translation agencies to help optimize the localization workflow and, as in the case of multilingual content management systems, to speed up the actualization of multilingual documentation or websites. As the typology in Figure 5 shows, the automation of the process increases from right to left. The model also shows the overlap in terminology database, translation memory and, to a lesser extent, localization tools used by translators and localizers. The translation-memory and localization programs available do vary however with regard to the number of available features. Software used by freelance translators oftentimes offers only part of the functionality available to localizers. These customized applications have become known as "light" or "front-end" solutions.

Translation and Localization Technology			
Localizer / Productivity Tools		Translator / Knowledge Tools	
DTP Tools	Term Extractors, Term Bases	Term Bases (Glossaries)	Encyclopedias
Quality Assurance Tools	Translation Memories Back Ends	Translation Memories Front Ends	Dictionaries
Project Management Tools	Localization Tools Back Ends	Localization Tools Front Ends	Digital Archives
Workflow Systems			DIY Corpora
Content / Globalization Management Systems			Concordances
Machine Translation			Specialized Websites and Newsgroups
Internationalization			Multilingual Knowledge Management

Figure 5: A Typology of Translation and Localization Technology

While our distinction between translator tools and localizer tools serves a theoretical and didactic purpose, the processes involving these tools and their individual advantages and disadvantages cannot and should not be seen as separate. The close interconnection of translation as the transfer of knowledge across cultural and linguistic borders and translation as part of a larger business process must not be neglected.

A bit of localization theory

Both the translation courses and the tool courses are flanked by additional courses that offer a more seminar-style discussion of localization issues. This third element of my overall approach aims at applying existing translation studies approaches to the field of localization and developing new ones.

Theoretical considerations of the localization paradigm should also address the changes that the industrialization of the translation process bring about for the professional lives of translators and localizers. This should include a description and critical analysis of typical localization workflow patterns and a warning against the translational constraints resulting from the use of translation tools.
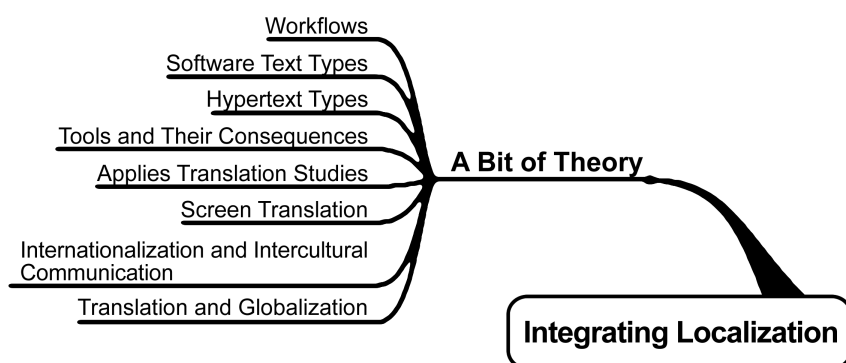

Figure 6: Integrating Localization—Possible Theoretical Issues

The main characteristics and advantages of translation-memory (TM) systems are widely known, and could be summed up as follows: Given the fact that technical documentation in general tends to be redundant, the use of translation-memory systems eliminates the need for repetitive translations of regularly recurrent textual segments. This refers to repetitions of the same or similar source text units within the same text (internal repetitions) or repetitions within a corpus of previously translated texts (external repetitions). The automatic recognition of previously translated segments

increases the stylistic, phraseological, and terminological consistency of the target texts, which constitutes a major quality improvement. The elimination of repetitive tasks leads to faster turn-around times, productivity increases, and lower costs, and at the same time frees the translator from time-consuming, boring, and error-prone tasks. Project management functions available within translation-memory tools provide, for example, statistical information about translated segments and thus allow for the better planning and monitoring of localization processes. Translation memories can be used over local or global networks, which speeds up team-based translation projects, and helps to secure consistency among translations produced at remote, yet interconnected sites.

Despite these undisputed advantages, translation-memory usage also includes a number of inconveniences, especially from the point of view of individual translators. Among the complaints from the translator community are the rigidity of source text structures, the dominance of the sentence or sub-sentence phrases as primary translation units, incompatibilities within one TM or between TM and term bases contents, faulty yet untouchable segments, the lack of creativity for the translator as autonomous text producer, the lack of co-text and context for the segments to be translated, and the lack of motivation or freedom to go beyond the simplistic source text structures and the preexisting translations imposed upon the translator by the TM system. Another problem with regard to the use of translation memories is the question of copyright and intellectual ownership of the translations that form part of the TM.

Given the dangers of a snowball effect of translation errors embedded in TMs, one must control the quality of segments stored (for source texts as well as for target texts) and the consistency of the content of TMs and term bases become essential for the overall quality of any translation project. Therefore, TM systems must provide for the easy manipulation and updating of existing TMs, including the automatic update and replacement of new or modified terminology. This quality maintenance is directly related with the reliability of a TM and thus with the quality of the work produced using a TM system. That sounds pretty simple, and all TM suites offer the necessary features for this kind of quality control.

The problem, though, is that the realities of modern, conveyor-belt-like localization projects, tight delivery deadlines and even tighter budgets mean that quality control of TM content is often not carried out thoroughly enough. As a result if this neglect, units stored in translation memories are often neither reliable nor consistent, which basically renders the main arguments for their usage obsolete.

In many case the use of TMs and other localizer tools thus leads to frustrated users. Many of the problems are caused by not seeing translation as an integral part of localization projects, and by not considering technology as and integral component of translating. Interesting, and I would add, rather

telling about some approaches to translation of the language industry, is a statement on the role of translators taken from a rather expensive-looking Trados brochure: "The translator or linguist is a language expert responsible for the creation of the translation as such. He (or she) focuses mainly on the content (of the translation) and not so much on the technologies involved or on the translation process (as a whole)" (Trados 5). I struggle to see why a company that is dependent on its TM sales needs to promote such an isolated, outdated, and utterly technocratic view of the translator and his or her doing.

This quote, however, seems to be symptomatic for an industrial system that creates a seemingly permanent frustration among freelance translators who feel either exploited and/or deprived of their linguistic and translational freedom or who "just don't give a damn."

Bob Clark (2003) has warned of the dangers connected with the strict hierarchy of the localization industry. He calls for the rehumanizing of translation, and describes rather well what is happening to individual translators within the localization industry.

So, in today's professional reality, home offices regularly convert into sweat shops with translators desperately trying to meet yet another unexpectedly advanced delivery date. Due to the size of the files to be dealt with and inhumane dead lines, many modern translators feel exploited and over-pressured. Yet, at the same time they are often times bored because of the monotony of their work, e.g. the translation of seemingly endless software strings. In addition, many typical localization text types such as resource files significantly cut into a translator's freedom, forcing him or her to-quite literally-count characters. In text types that due to their functions and structures would give translators a little more creative leeway beyond bilingual bean counting, the dictatorship of terminology presets (many of them established by linguistically-challenged software developers) and the sacrosanctity of translation memories restrict the hermeneutical activities of translators right form the start. Just like that, the advent of translation memories and the "one-size-fits-all approach" they represent, have effectively reintroduced "the phrase" to the throne of translation units.

Some of the central components of modern translation project management contribute to a translation reality that is in many ways diametrically opposed to key paradigms of modern translation theory. What room is there, for example, for Hönig and Kußmaul's "degree of differentiation" (1982:58-63), which allows, or better, implores the translator to add information, to leave out information, to alter the text where necessary? How many times are users of translation memories faced with a couple of source-text sentences that would sound just lovely if made into one in the target language? Of course, technically, that could be done by changing the segment alignment. But how many translators would do this, and how many project managers or clients would accept it? Under these circumstances, can

translation still be "the creative give-and-take of intuition and cognition" that Paul Kußmaul writes about (1998:49 and 2000)? To be fair, however, with regard to many typical software types, e.g. on-screen texts and strings, this kind of approach would represent a theoretical overkill. But does the same hold true for manuals, for e-learning material, for marketing texts? And what about instruments of coherence? Anaphora, cataphora, isotopy, paraphrasing, substituting? Forget them. In a text that is a "just-in-time document", that is less a text than a momentary assemblage of content fragments within which every fragment, every phrase can become the readers entry point, "repeating" becomes the one and only resumption strategy. Think that's bad? Wait for content-management systems.

My reason for stressing the negative impact on individual translator is also to sound a warning. Translating within larger localization projects or for the pitiful word fees of many agencies can no longer be advertised as an attractive and challenging profession (not to mention it being a lucrative one). Many excellent graduates of translation schools are already migrating to new, more rewarding professional fields. And those working in the localization setting are constantly looking for ways up or out, making technical and software translation more and more and entry level job or a way to survive financially until something better comes up. The results of this are a lack of qualified and motivated beginners, and a translation brain drain, i.e. the professional escaping of qualified and experience translators.

Localization and, above all, internationalization can benefit from Translation Studies. Anthony Pym (2003) has mentioned some of the possible links. Finding ways of applying Translation Studies to localization (and developing new approaches) will be an important challenge for academics in the field. A comparison of subtitling approaches and software and website localization, for example, will show interesting similarities between these two types of screen translation.

We should also become more involved in thinking about the ways in which translation is related to computer-mediated intercultural communication, and how it fits into the workflow of localization processes. In that regard, it would be interesting to see if, for example, consistent and resolute post-alignment of thoroughly researched and revised translations could not lead to higher productivity, better quality, and more consistency in TM usage.

Furthermore, scholars might want to look at and compare text types involved in website and software localization. Using a typology of software text types (see figure 7 for a simple representation), analyses could focus on the textual characteristics, inter- and intracultural differences or technical constraints of these specific texts, which would be one way of preparing future translators (and technical writers) for the advent of XML-based globalization management systems.
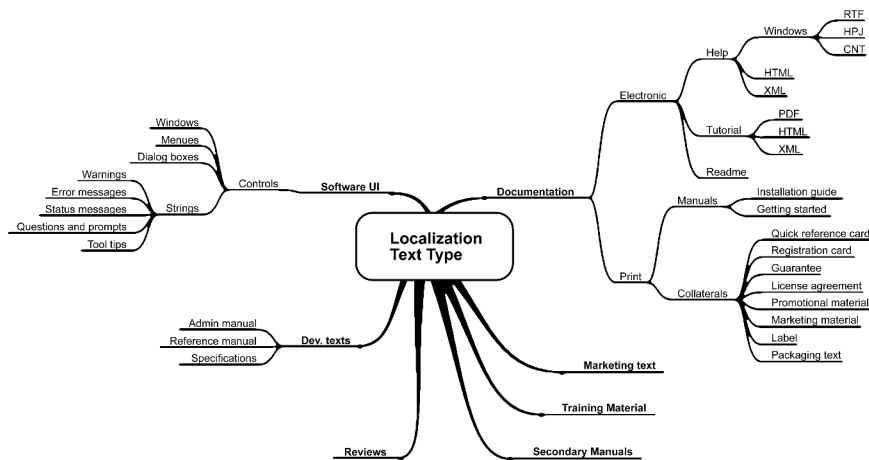
Figure 7: A Typology of Software Texts

Conclusion

The model described above calls for a scalable approach to integrating localization into translator training program. The approach reflects the institutional constraints of a traditional four-year program. The model has proven to be flexible enough to allow students interested in technical translation, translation technology and localization to combine numerous obligatory and optional courses for a specialization in this field within the traditional *Diplom* program. The obligatory four-level module on software translations (English to German) guarantees the student's exposure to the dominant text types and tools involved in software and website localization. The optional second part of the overall approach, the courses on electronic tools for translators, allows for an individualization of the learning pace by letting students select the courses on the basis of their prior experience. The courses on tools are also very well suited for conversion into e-learning units. The third component of the approach allows for a more thorough and critical analysis of the localization paradigm. Students can write term papers and their final theses on the issues mentioned above and might even go on writing their dissertation about Translation Studies and localization.

The courses offered within the above model can be easily combined with other translation courses, for instance in order to cover other relevant language combinations like Spanish and German, or English and Spanish. They could also serve as the basis for more technology-oriented program on translation tools and localization project management.

References

Clark, Bob. 2003. "Are Translators Getting Their Voice Back?". *The LISA Newsletter. Globalization Insider* 12. 2.4.

Hönig, Hans G. and Paul Kußmaul. 1982. *Strategie der Übersetzung. Ein Lehr- und Arbeitsbuch*. Tübingen: Narr.

Kußmaul, Paul. 1998. "Die Erforschung von Übersetzungsprozessen: Resultate und Desiderate". *Lebende Sprachen* 43. 49–53.

Kußmaul, Paul. 2000. *Kreatives Übersetzen*. Tübingen: Stauffenburg Verlag.

Pym, Anthony. 2003. "What Localization Models Can Learn from Translation Theory". *The LISA Newsletter. Globalization Insider* 12 2/4.