

The End of Translation as we Know it

An industry
watcher speculates
on translation
post-Y2K

by Bert Esselink



Bert Esselink

The number of articles written in the last months of the millennium must be enormous. I could have called this article "Translation in the New Millennium," but this title would have been far too pretentious. Nobody knows how translation and translation technology will develop over the next few years. However, it is obvious that there will be many changes. The so-called New Economy, which is so heavily characterized by globalization, will also have an impact on translation technology and the translation or localization industry.

Even though the art of technical translation will always remain the same, there will undoubtedly be changes in process workflows and the way in which translation is managed. The next few years will show us the end of translation as we know it. Below you will find a quick retrospective, and some possible scenarios in which translation technology and the industry might develop.

Localization Models

In the beginning of the '90s, the localization industry was finding a place for itself somewhere between the software industry and the translation world. Many companies emerged that were—or claimed to be—specialists in managing,

translating, and engineering complicated software products. After years of mergers and acquisitions, we are now left with an industry consolidated into a dozen MLVs—multilanguage vendors with an international presence—and thousands of smaller SLVs, the single-language vendors

**After years of mergers
and acquisitions, we are
now left with an industry
consolidated into a dozen MLVs,
multilanguage vendors with an
international presence, and
thousands of smaller SLVs.**

or traditional technical translation agencies.

Traditionally, localization projects didn't start until the entire source product—usually US English—was finalized and had shipped. Customers in non-English-speaking countries often had to wait for several months before they could purchase the product in their native language. When the software and localization industry matured and globalized, product lifecycles, time-to-market, and simultaneous shipment of different language versions

became important issues for software developers. Publishers wanted to ship all language versions either simultaneously or within an acceptable period after release of the English product. This automatically resulted in a different localization model, where translation started while the English product was still being developed. This new, parallel model created the need for translation leveraging tools, such as translation-memory and software-localization tools, to enable localization vendors to re-use translations with each intermediate product update or release.

Now, it seems that this model will also be outdated soon. The main reason being that localization *projects* will soon change into continuous localization *flows*, i.e., steady streams of updates and product revisions. For example, in the next few years, more and more software applications will run from servers, mainly through the Internet. Static software products shipped in boxes will gradually disappear, as will localization projects with a definite beginning and end. Besides, the frequency of Web-site updates will become an important competitive advantage. For example, if a Web site is published in six languages, each change made in the English version will have to be made in the other five languages within days, or in the case of important news announcements, within hours of the English change. As a direct consequence, translation and localization will be processes that run parallel with development and content creation.

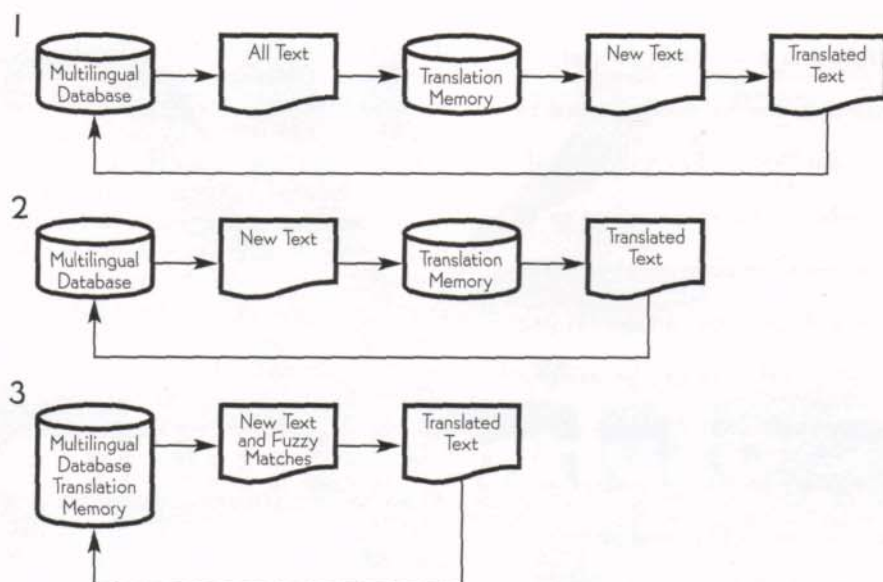
The only way to keep frequently changing and updating material manageable is to store all the information in a database format and content-management system. Not only documentation, but also software, online help, and Web-site text and graphics will be created, managed, and published using database technology. Today, most e-commerce Web sites selling products online are already based on giant databases containing product information that generate Web pages on the fly using scripting languages and active server pages. To efficiently and quickly translate these so-called dynamic or live Web sites, advanced translation technology needs to be utilized.

Translation Technology

With the change of general software and documentation creation and publishing technologies, translation technology will also have to be adjusted. Because the translation turnaround time will become the key ingredient for industry competitiveness, the number of steps between content creation and translation needs to be reduced to a minimum. Ideally, translators should sit at the same desk as the software developers and technical authors and immediately translate every single written word. The closest we can get to this is to grant translators or translation agencies access to the source-information database, to enable them to directly translate batches of new material as soon as it is released or approved.

Changes or additions to the database should automatically be tracked and transferred to the translators. A real-life scenario would be that technical authors add a new piece of text to a Web-publishing database,

Ideally, translators should sit at the same desk as the software developers and technical authors and immediately translate every single written word. The closest we can get to this is to grant translators or translation agencies access to the source-information database, to enable them to directly translate batches of new material as soon as it is released or approved.



Three Translation Technology Models:

- 1) All information is extracted from the database, processed by TM so only new text is translated.
- 2) Only new or changed information is extracted from the database and processed by TM.
- 3) The content-management database contains a TM engine that pretranslates all new or changed content so translators will only receive new text or fuzzy matches.

and then indicate that there's new information to be translated in a specified set of languages. The new text is automatically pushed to previously selected translators or translation agencies, who find this new project in their translation inbox. After translating the text using a translation-memory system, the translated text is automatically transferred to the correct language section of the multilingual database for validation, QA, and publication.

In this setup, there will be two databases running simultaneously:

- A *multilingual* database containing all source-language text that is written and all language versions of the Web site, online help system, documentation set, or even software application.
- A *bilingual* translation memory database maintained by the translator or translation agency.

A much more efficient solution would be to integrate a translation-memory engine in this large, multilingual database that contains both source and translated material. This engine would automatically analyze all new content created, and pre-translate text using full or fuzzy matching before it is sent out for translation. The translation-memory engine could run parallel with a structured authoring/controlled English tool which guarantees

maximum leverage, both in the source language and in translations. If structured authoring is used, and the language pairs are feasible, the multilingual database could also be linked up to a machine translation engine to pretranslate material that wasn't found in the translation memory.

In this scenario, translators would only need a translation editor or limited version of the translation-memory engine, which for example only processes internal matches. Translators eventually should be able to maintain real-time translation-memory connections with an Oracle or SQL database through TCP/IP to enter their translations. The same database should be accessible for reviewers, both from the translation agency and from the customer to avoid delays caused by file transfers. Especially with a fragmented translation model, a final QA of all translated material before it is published is a must.

Localization Industry

What would be the consequence of total *database-driven structured translation* for the localization industry? Most importantly, there will be no more fixed projects with a definite beginning and end. The main advantage of this approach for translation vendors will be workload spreading. Trans-

Another advantage is that translators can be easily linked to certain publishers. For example, if a particular software company or e-commerce firm updates their Web site on a daily basis, a translator or group of translators can be assigned jointly by the translation agency and publisher to always translate these updates. This will ensure a consistent style and maximum company and product knowledge by these translators. The number of translators required can be directly linked to the average throughput generated by technical authors or developers.

lators really work side-by-side with the developers and authors. Another advantage is that translators can be easily linked to certain publishers. For example, if a particular software company or e-commerce firm updates their Web site on a daily basis, a translator or group of translators can be assigned jointly by the translation agency and publisher to always translate these updates. This will ensure a consistent style and maximum company and product knowledge by these translators. The number of translators required can be directly linked to the average throughput generated by technical authors or developers.

A disadvantage will be that in the case of many small updates or additions, translators would likely be working on several small projects every day. For example, if publishers generate an average of 600 translatable words per day, one translator will need at least three projects from each publisher to generate sufficient revenue. Leverage-tracking systems will monitor the number of words that are sent out for translation or review, so the work can be invoiced at agreed times. Establishing long-time relationships and contracts between publishers and vendors will be essential for the success of these automated translation workflows.

In this scenario, even the project management and file distribution performed by translation or localization agencies would be a delaying factor. The information should be directly exchanged between publishers and translators. Where does that leave the agencies, or the large MLVs? First of all, translation quality needs to be verified. Even though translation can be outsourced to freelance translators or third-party agencies, most large publishers will have to depend on vendors to guarantee the overall quality of the delivered translations, to manage terminology, to manage the translation resources, and to select and assess the resources used.

Also, multilingual database publishing only works for technical content that can be translated almost literally. It doesn't apply to true localization, i.e., the creation of local content for different countries. For example, most companies will want to adjust the home pages of their Web sites to a local situation, containing regional information such as marketing information, local offers, and regional office information. So MLVs will need to have the expertise and resources to truly localize, or basically rewrite information in as many different languages as possible.

Structured Translation

Companies that want to go global in the most efficient and cost-effective way pos-

sible cannot ignore the power of multilingual database publishing. Structured translation will go hand in hand with structured content creation. In a few years, content will be automatically replicated from a central database to locations all over the world, translations will be automatically transferred back into the central repository, and localization vendors will find themselves managing people and processes instead of temporary projects.

The industry and translation technology still have a long way to go in the new millen... next few years.

Bert Esselink has been active in localization since 1990. After graduating in technical translation and doing coursework in programming and computational linguistics, he worked for several years as a software translator, localization engineer, and technical manager. In 1996 he joined ALPNET in Amsterdam as localization manager, training new engineers and localization specialists, and coordinating software localization projects. Since early 1999 he has taken on a new assignment as corporate globalization manager for ALPNET. His first book, A Practical Guide to Software Localization, was published in 1998. Bert can be reached at bert.esselink@planet.nl

