What is Language Technology?
Chris Langewis

Comparing Basic Features of TM Tools
Angelika Zerfaß

Technology and the Freelance Translator
Jonathan T. Hine, Jr.

Finding the Right Language Technology Tools
David Shadbolt
This guide is published as a supplement to MultiLingual Computing & Technology, the magazine about language technology, Web globalization and international software development. To subscribe to MultiLingual Computing & Technology, go to www.multilingual.com/subscribe or send your name and address to MultiLingual Computing, Inc.

Publisher / Editor: Donna Parrish
Editor-in-Chief: Seth Thomas Schneider
Associate Publisher: Jennifer Del Carlo
Managing Editor: Laurel Wagers
Associate Editor: Jim Healey

Editorial Board
J eff Allen  
Henri Broekmante  
Bill Hall  
Andres Heuberger  
Chris Langewis  
Ken Lunde  
John O’Conner  
Mandy Pet  
Chris Pratley  
Reinhard Schäler

Staff
Becky Bennett  
Sandy Compton  
Kendra Gray  
Olya Helms  
Doug Jones  
J erry Luther  
Courtney McSherry  
Bonnie Morrell  
Sarah Ragan  
David Shadbolt  
Aric Spence  
Cecilia Spence  
Kevin Watson

This guide is published as a supplement to MultiLingual Computing & Technology, the magazine about language technology, Web globalization and international software development. To subscribe to MultiLingual Computing & Technology, go to www.multilingual.com/subscribe or send your name and address to MultiLingual Computing, Inc.

Subscription rates:
One year (eight issues) $58
Two years (sixteen issues) $100
Outside the United States, add $20 per year for postage
Language Technology

I am not a very handy person. So, when I walk into a hardware store, I am amazed at the choice of tools, implements and machines that are available. The usefulness of some of the tools is inherently obvious. We can all appreciate a fine well-crafted hammer and imagine how it can make a carpenter’s job easier. Some of the things I find there are a total enigma to me. What does it do? Why would I want it to do that? And, then, once I finally understand its function, what clever person designed this?

I think this must be what it is like to approach the choices of language technology products. The range of capabilities and applications is mind-boggling. So, we’ve put together this supplement to introduce you to the field of multilingual language technology.

Chris Langewis gives us an introduction to language technology with definitions of all the types of tools and descriptions of how they have evolved to make the language professional’s job easier. Translation memory products are some of the most useful tools, and Angelika Zerfaß describes their common features. Jonathan Hine walks us through the technology used by a typical freelance translator and how that translator makes the tools work. So, once you have a grasp of this technology, what companies produce what? David Shadbolt introduces us to some of the producers in the industry. We follow that with a table of product listings to give you an idea of what is out there. We gathered much more information than is in the table — as you can see at www.multilingual.com/ltDetail

It might be helpful to view the field of products as the parts forming a modern transportation bridge. The foundation is made up of engineering tools and computer resources — font libraries, character sets, internationalization and localization routines that help form the basis of translation and localization processes. The horizontal girders are comprised of translation software such as dictionaries, translation memory databases, machine translation products or desktop suites of these tools. All of these help to form the road that carries localized software, e-mail, data, documents and other translated materials. The bridge’s suspension structure represents the emerging array of multilingual workflow systems, including content management, project management and workflow monitoring. Together these products are changing the nature of communications by creating a high-speed bridge across the multilingual gulf.

There is a Chinese proverb: “To do good work, one must first have good tools.” There are some fine, well-crafted language technology products available. It’s time to pick up the ones you need and do some good work! — Donna Parrish, Publisher

CHRIS LANGEWIS is vice president for marketing at Logos Corporation, teaches computer-assisted translation at the Monterey Institute of International Studies and is a member of the MultiLingual Computing & Technology editorial board. He can be reached at chris@langewis.com

ANGELIKA ZERFAß is a freelance consultant and trainer for translation tools and localization-related processes. She can be reached at azerfass@debitel.net

JONATHAN T. HINE, JR. is a freelance translator and also teaches technical translation and business organization for language mediation professionals. He can be reached at hine@cstone.net

DAVID SHADBOLT is a research editor for MultiLingual Computing & Technology and can be reached at david@multilingual.com
I remember a conversation at a major business conference in the 1980s. I was sitting with a group of executives from Europe, Asia and the United States. We were discussing the problems associated with language and translation of our various high-tech product lines across the various languages of our markets. One opinion that surfaced caught me by surprise — that the whole issue of translation in our community would become academic since English would become the universal business language anyhow. A lively discussion ensued, needless to say. We concluded that even though many people in many countries could speak English, they did not necessarily wish to use that foreign language in their daily routines. If you are to successfully sell your product in a foreign market, it should be adapted to the local user environment.

History has since then confirmed that our language, localization and translation issues have not only stayed with us, but have dramatically increased in some critical aspects. Some of the phenomena conspire to seriously challenge today’s multinational marketer. The products that we market and the format of their translatable content have become much more complex. The time-to-market requirements have shortened dramatically. The volume of the content that requires translation has grown exponentially. The subject matters of the translatable content is expanding. The time to market is reducing. Technology developments mean that many companies are now operating in a large number of foreign markets, some of whose languages are considered to be “exotic.”

So, we are faced with localizing a growing number of increasingly large, complex products targeted to more exotic foreign markets, and it has to happen faster. Sounds like a challenging situation, which portends to strain both workforce and budget. In response to these issues, lots of energy and thought have gone into the idea that a technological solution might exist to resolve this dilemma, inventing thereby the field of “Language Technology.” One definition of this is a computer-based tool or program which supports humans who are involved in adapting a product and its language content for optimum usability in a foreign market.

Many products and “solutions” have been developed that are designed to address the problems of getting your products into foreign markets more quickly and elegantly. But which product do you choose? How do you integrate these “solutions” into your business processes? How, exactly, does one address an issue as complex as a product in multiple alternate languages? This content could be manifested as a printed page, a Web site, the user interface of a software product, on-line material or any number of other representations. The problem is not only that the translation of the content is a complex issue, but also the process of re-publishing the product in each alternate language is complex as well. In fact, if you start with a product in one language and re-publish it in another language, only 25% or so of this entire process is actually language work. Translation. The overall process consists of many steps, including content extraction, project and workflow management, translation, product reconstruction, review, QA and testing.

Where and how do we apply language technology? The trick is to identify the critical path in the overall process and to isolate the necessary human tasks versus those tasks that can be handled by technology. We can divide the field into categories, even though there are overlaps where a given tool might appear more than once. But generally the categories are translation tools, language processing resources and localization tools and multilingual workflow systems.

Translation

Translation is the process whereby the meaning, style and information of a piece of written text in one human language are reproduced in another written human language.

In the “old days” a translator would receive a stack of paper containing text of some kind, and would be asked to return a stack of paper with that text in a different language. A typewriter might be used, but it might also have been a pen or pencil. In the process of translating, a dictionary and thesaurus would be used, and the translator conferred with colleagues and document authors to resolve terminology and “meaning” issues. Unique terminology (specific to the subject matter of the assignment) would be captured bilingually on file cards for future reference and to ensure consistency of use in the document being translated. The whole process required static text and took a long time.

This process worked for centuries. Images of monks using quill pens by candlelight, bent over scrolls of paper, translating the Bible comes to mind. They had time on their hands. The source documents were fairly static — very few updates were issued (presumably). But modern times are more difficult. Computers with word processors came along. Documents proliferated, and translations of them were needed rapidly. Translators now needed computers and WP systems to keep up with demand. One very critical fact was learned in all this — that a person’s output of translated material is basically the same whether you type on a computer or type on a typewriter or write with a pencil. Generally, 2,000 words per day is the amount of productivity you can expect, with variations depending upon language, subject matter and complexity. This fact dictates much of the critical path in organizing projects.

To complicate matters, documents could change during the translation process. When a change was made in a source document (an update) in mid-process, it wreaked havoc. Do you cut and paste the new material into your translation? Where are all the differences? Just finding the changes is very time consuming and error prone. So, let’s invent some language technology to address this and other issues.

Along comes the first wave of translation tools, one of the main categories within the language technology field.

Translation Tools

Translation tool is the label given to that category of computer program that supports humans who are performing translation tasks.
Key Benefits

Distributed Environment
Your translators can work from multiple locations anywhere in the world and can share the same memory. Translation memories can also be located anywhere in the world.

Flexibility
There is no restriction on TM software used by your translator. You have a wider choice of translators for a project.

Ease of use
No retraining necessary! Translators work in MS Word.

Management
There is a significant reduction of project managers' technical and preparation tasks.

Quality
Translations will benefit from increased leveraging and consistency.

Control & Security
Your translation capital is protected. Your TM is stored safely on your disks, so there is no need to distribute them.

T-Remote Memory

Next Generation Translation Memory Systems

- Run your translation team in real time and share the same data.
- T-Remote Memory is powered by the Internet and allows you to run several translation memories, terminology databases and MT output simultaneously!
- Translators work from remote locations in real time, and project managers manage just a single database.

See http://www.telelingua.com for more information.
The tools merely provide environments which enhance the ability of humans to provide timely translations of high quality in a cost-effective manner. Translation memory (or sentence memory), terminology management systems, bilingual editing systems and machine translation are examples of this type of tool. Interestingly, these tools can have dramatic impact on how you approach and conduct everything from product development to multilingual publishing.

Translation Memory. Translation memory (or sentence memory) is the label given to that category of computer program that stores sentences, along with the translation of each, in a database.

The single most important and useful tool in this category is translation memory (TM). This tool is quite simple in principle, yet can become very sophisticated in practice. TM is a database into which a translator places sentences from a source document being translated and then correlates corresponding translated sentences (target sentences) with each source sentence. When the document translation is finished, you have a database which correlates the entire source and target document sentence by sentence. However, this does not get the project done any faster. You can still get only 2,000 words per day of translations created. But several extremely powerful benefits are endowed upon future processes. You may be able to get the job done sooner, and you might save a lot of time and money.

Since we can’t materially change translator’s productivity (words per day), we must extract increased productivity by insulating the translator from process-related issues. We do this by finding ways to invoke the Three Golden Rules of translator productivity: 1) Deal with fewer words. 2) Avoid repetition. 3) Avoid non-translation-related tasks. In fact, to qualify as a translation tool, one of these rules must be invoked. If you can’t relate the effect of a given tool to one of the Golden Rules, it won’t be of any help to translation productivity. To see how TM fits these criteria, here are a couple of examples.

Let’s assume your translator is partly done with the translation of a project (which could be a document, a Web page, strings from a resource file or any other content). The content owner decides that some changes are needed in the source material and delivers an updated source file to the translator. This is traditionally the worst possible scenario. Time will be wasted either by starting over, or by researching changes, cutting, pasting and so on. These activities are repetitive. They increase the word count being dealt with, and they involve non-translation work.

But the project has been done using TM. We simply stop work on the original document and start with the new version of it. The TM system will look at each sentence of the new document and compare it to what’s in the database. If it finds an exactly matching source sentence, then its previous translation is still correct, so it is used in the “new” document’s translation automatically (leveraged) without translator intervention. If a sentence in the new document is almost the same as the original (a “fuzzy match”), the original translation is reused with a note to the translator to check and edit it to reflect the changes. Any new content is simply passed through to be translated. In this manner, the update causes almost no disruption to the translator’s productivity. All prior work is recycled into the updated project. All of our Golden Rules are actively applied: no translation work was repeated, very little non-translation related work was done, and the translator could focus on translating only “new” words.

The power of this capability lies in the effect it has on overall process options. The ability to deal eagerly with updates enables you to begin a translation project on unfinished content, thereby starting sooner and finishing sooner, critical in managing time-to-market of critical materials. Further, translating future releases of similar materials may be greatly simplified. Any content in a new release that is the same or similar to an older version will enable TM to leverage its translation, thereby saving much time and money. Also, if you have products that share pieces of content, you need only to translate this content once and then leverage that translation into each recurring instance.

Some TM systems come with a built-in editing environment so that translators view and translate content within the TM session. Other TM systems use third-party editors such as Microsoft Word to view and edit, but then capture and catalog the results in a database for you. In any case, the ability to view and edit bilingual is critical to terminology management.

Terminology Management (Glossary or Dictionary). Terminology Management is the label given to programs that catalog words and phrases along with pertinent related information in a database in a manner conducive for use in linguistic applications.

Another powerful feature that many TM systems employ is a Terminology Manager. Using terminology consistently within a given project is critical to achieving a high quality of translation. In going from one human language to another, any given word might have multiple meanings and therefore will have multiple possible translations. For example, in English the word pole might mean stick or rod, but it might also refer to one end of a magnet or to the top or bottom of the planet Earth. Each of these would require a different translation depending upon context. Also, new words are invented which require creative translation that must be applied consistently, not only within the project at hand, but subsequently among similar projects as well.

Instead of relying on note cards to catalogue these details, the terminology management software can be used to create glossaries. Each word with unique usage in your particular content can be placed in a glossary, be assigned a subject matter, translation, definition, and other characteristics. The glossary interacts with TM so that translators are reminded of the proper translation for terms that appear in the content being worked on. This helps to create consistency and quality within a given project, among a series of projects, and within a workgroup translating in parallel. The content owner can archive the glossaries and reuse them in future works, thereby saving a great deal of time and money by avoiding rework.

Machine Translation. Machine Translation (MT) is the label given to that category of computer program that converts a sentence of human language into an equivalent sentence in another human language.

This writer teaches a translation technology course (Computer Assisted Translation) at the Monterey Institute of International Studies in Monterey, California. At the beginning of each semester, I ask the class to raise their hands if they believe that MT works. Typically, no hands go up. After all, only humans can really translate. Why else are they at the Institute to earn a degree in the subject? However, by the end of the course, all hands go up in response to the same question. The difference lies in knowing when, why and how to apply the technology.

MT has the appearance of producing flawed, even laughable translations. However, well deserved this reputation, there is ample reason to believe otherwise. MT is like an automobile in that it is the driver that determines where it goes and whether the trip is successful. By understanding the critical path to maximizing results, one gets radically different results from the same car or from an MT system.

Without going into great detail, it’s relevant to know what makes MT “tick.” The basic idea is to create a translation of a whole sentence automatically. The various products that exist accomplish this somewhat differently, but the general pattern is to parse the sentence into words, look the words up in a custom dictionary and establish possible roles for each word (noun, verb and so on); based upon grammatic and syntactic rules of the
The factors critical to making this process work properly are 1) correctness and clarity of source sentence; 2) completeness and correctness of dictionaries; and 3) sufficiency of rulebase (grammar, syntax) that is being used. Given a perfect situation in the three factors above, an MT system can very likely produce perfect translation. But any small deficiency in any of these factors will result in a very poor translation. Using colloquial expressions or slang will confuse most systems, for example. If a word is found to be missing from the dictionary (a new technical term, for example), its role and meaning cannot be established, so a very bad translation will result. Fragments of sentences are difficult for MT systems to parse because the rules of grammar generally don’t apply. Human skills are required to comprehend them. Worst of all is ambiguity. What is obvious to a human may be confusing to a machine. For example, the sentence “the boy threw the rock at the window and it broke” is clear to a human who understands fragility. A machine may need to guess whether the window or the rock broke. Therefore, to maximize results you know what needs to be done ensure that all input sentences are grammatically correct and unambiguous so that all words in the sources are properly represented in dictionaries, and adjust the rulebase to cover any special writing style of the source content.

Even given the optimum scenario of perfect inputs and preparation, you still need to review and “post-edit” the output to ensure quality, which can be a significant task in itself. Granted, all this can be difficult to accomplish. It will take significant effort to glean any reasonable results from an MT system. So, when is it useful? A project might require turnaround times that are not possible without MT. There may be times when a project is just too large for humans, such as 500,000 pages of technical manuals. When optimized, MT is very fast. A high-quality draft can be produced in minutes, leaving only the QA process. In short, large projects with short turnaround requirements are good candidates. Also, special projects that can accept draft or gist translation, where less-than-perfect is acceptable, are good candidates.
Some additional ideas include gist translation of e-mails, translation of "newswire" content on the Internet, and on-line catalog item descriptions in fast-turnover sites. Generally, where speed is mission critical or volume is high, MT may offer a solution. Creative application is always needed, but can be rewarding.

After the Translation

Regardless of how we got there, let's assume we have our content translated. We now have the content of our Web site, document or software product in multiple languages. Even though the translation has been completed, much work actually remains to be done. We need to put the content in a "live" environment to make sure it looks and reads right and to ensure that the resulting translated object functions as the original did. Also, we need to ensure that the resulting translated product is what will be considered "native language" quality in the target market. This is especially critical with software and Web-site projects. We're very likely to find some problems. These problems that come to light as a result of translation will require some engineering resources, such as localization or internationalization, to resolve. Fortunately, a category of language technology has evolved to help us here.

Engineering Resources

Engineering Resources are those tools and capabilities that are used in conjunction with translation activities to adapt a domestic product to a foreign market environment so that it has the same technical suitability as a product that has been created in that foreign market.

Localization and Internationalization

Localization is that process whereby an object such as a computer program, multimedia presentation or document is not only translated, but also is adapted to another culture or foreign locale. Such things as character set handling, date formats, metric/English measures, comma/decimal conventions and other elements are considered in addition to the normal language issues.

Internationalization is that process whereby an object is designed and developed in such a manner to minimize or eliminate the need to perform specific localization tasks after translation.

A good example of a variety of localization and internationalization issues is found in software that is marketed in multiple countries or locales. We have a combination of program code and textual content. The content shows up in menus, dialogs, on-line help systems and more.

Content Extraction. The first problem that arises is how to organize the content so that translators can deal with it. In the "old days," content (menus, error messages and so on) was sprinkled throughout the program code wherever the programmer decided to place it. You needed a special "extraction tool" just to get to it and to put the translated text back where it belonged. Most of these tools were custom made by the software developers. Very rapidly, the cost and awkwardness of this process led to the first "internationalization" strategy — put all the content in a file and have the program refer to it and call it up as needed. So, the Resource File was created. Today, almost all software has content in separate, easy-to-read files. Simply translate this file or files, preserve its format, and the program displays the foreign language content during program execution. Without the resource files, you'll need to build or buy extraction tools.

Coincidentally, TM tools perform many extraction tasks for you. In fact, you might choose a particular TM tool because it is able to deal with the particular resource structure or text file that you have. For example, today's content is often held in HTML, XML or other sophisticated formats. Most TM systems can read and regenerate these structures and thereby insulate the translator from the complexities of dealing with the embedded coding. In TM parlance, these extraction devices are called filters. Before selecting any tool used in a language technology context, make sure that it has a filter for your particular file formats.

Character Set Handling. In most modern software and Web-site environments, foreign character sets are handled at the operating system level or by the browser. You might still find some issues, however. If you start with English or any language that uses a single-byte character set, and the translation is into a double-byte set (such as Japanese or Chinese), testing is a must. Even more problematic are the right-to-left (Arabic) and bidirectional scripts. Your program code may be fooled by some of the characters' encoding. In the case of Web sites that hold content in a database, the exotic characters can foil lookup routines and string handling code. It's wise to anticipate these factors.

A "code sweep" may be in order — where a special tool scans your program code to identify telltale areas where character encoding will trip you up. Many developers have created custom tools to help with this, but many are available on the market as well. Of course, as developers get experience with these factors, they anticipate and internationalize their code by avoiding the pitfalls to begin with.

Another very effective way to check your Web site, program, or any other object that you will translate is to pre-translate it and run it through QA. This can be done very inexpensively by running all of the content through an MT system, available on-line for most languages at little or no cost. Of course, you can always purchase a system for local use at a variety of prices. In any case, you're not concerned so much with the quality of the output, just with the character set and other factors unique to the language. If your software works properly with MT content, it should be fine with human-translated content as well. Why wait until it's too late to find out that a problem exists?

Sizing Issues. One localization problem that often arises is related to size. As compared to English especially, most other languages use more space to say the same thing. Going into European languages may require 30% or 35% more room in a page, dialog box or menu. So, even if you have a perfect translation, it may look corrupted when displayed in a dialog box that's too small. You need to adjust the software's program code to change the dimension of dialog boxes or to change the size of a button that has text on it. Also, as the spelling of items in menus changes, you may need to rearrange the list or to assign new hot-key sequences. All of these activities typically require special training. If left to translators alone, it may prove to be very inefficient, yet it can also be very inconvenient to tie up a development team with the tasks.

Localization Tools. These problems have been so pervasive that special localization tools have emerged to deal with them. These tools have many traits in common, while each has special features as well. Basically these tools are similar to TM in that they have filters that enable you to read and extract content from a variety of filetypes, and they use a database of some sort to house content and its translation. As with standard TM functionality, you can leverage translations between updates as you work and between versions over time.

What differentiates these tools is their ability to operate in a WYSIWYG mode — the system shows you the content that requires translation in context (like a dialog box or menu) and enables you to translate it. At the same time, you can easily adjust button and dialog box dimensions and configurations to suit the new language size requirements.

An innovation on the part of some localization products is that you do not need to
have the source code or even the developer’s resource files to be able to create a completely localized version of the product. You can operate directly on the executable version. Also, at least one of the localization products can do an automatic pseudo-translation of your software product or Web site. Source-language text is partly replaced with characters unique to your target language and strings are expanded to mimic a translated effect. This way you can test run a product and see if any character set tolerance issues exist or if any reconfiguration is needed.

Given a perfect scenario, developers internationalize a product so that it will operate correctly regardless of the language that it’s translated into or the locale in which it operates. What the developers miss remains as a localization task after the translation is done. Fortunately, we have tools to help us here.

So, we see that we can take either a whole product or the language content of a product and get help from language technology in getting it translated and localized. But earlier it was noted that this is only a fraction of the entire process that ultimately yields a product in a foreign market. Organizations with multinational experience have learned that much of the effort and cost is hidden in process management. This has caused the emergence of yet another category of language technology — multilingual workflow systems.

**Multilingual Workflow System**

Multilingual Workflow System is the label given to that category of computer program which creates an environment that supports and orchestrates a range of activities that facilitates the development of multilingual products.

The first time an organization decides to convert a product (or translate a Web site, for example) into a foreign language version, I call it a Stage A organization. The translation is handled much like an event. You finish the source language version, freeze it, market it domestically and then give it to someone to localize. Some months later you’re ready to go to your foreign market. By the time you get rolling, a new source version is ready, thereby obsoleting your translated version. Oops. To avoid multilingual versionitis, we need to develop a way to release the product in multiple markets simultaneously by foreseeing the leadtimes and the necessary activities. Let’s use some language technology to increase productivity and set the stage to leverage from one version of product to the next. This would make you a Stage B organization, where you actively anticipate foreign market requirements.

You’re successful with your overseas marketing. You are now global; the planet is your market. When you design product, you have all of your markets in mind, and you design to local requirements, incorporating feedback from the field. You publish your products multilingually rather than in one language and then in other languages. You internationalize your products to minimize localization issues that emerge from the translation process. You’re now a Stage C organization. But you learn that managing the overall process is getting very cumbersome and difficult. Many products, many languages, several translation vendors, many groups and individuals within your organization share ownership of the overall picture. What now?

Fortunately, there is some light on the horizon. A new category of language technology is emerging — the multilingual workflow system. Implementing this would put you at Stage D. There are several possible incarnations of the technology but the objective is similar with all, which is to smooth the process through...
the pipeline from product development to the final delivery of localized product.

The Process

Even though we can isolate individual translation and localization tools and discuss their benefits and features, that’s still just the tip of the iceberg. It’s critical that these tools and technologies are applied to create efficiency, but it’s also necessary to make sure that they fit into the greater context of your global plans and infrastructure.

A sophisticated process flows all the way from product planning to global delivery and would contain the following elements: define global product needs; assess cost/benefit per language; design multilingual publication strategy and process; design product with internationalization; identify translation/localization resources; deliver translatable to global manager; deliver files to translation/localization resources; translate/ localize; manage and distribute updates to work in progress; get and disseminate status information; manage feedback loop to product development and marketing; retrieve content files; build product to test; test — feedback loop; QA — feedback loop; and publish/deliver.

The pacing of items in managing this process is often hidden in the logistics of file transfer and control, and in managing the great variety of resources involved in the process. The most difficult thing is often being able to get the information you need about status. There are many parallel activities, and there is a mix of in-house and outside resources involved. Disparate groups need to be closely coordinated. What may happen in a complete-ly manual system is that a file may sit in some one’s e-mail box passively for an amount of time, may be misplaced or misdirected or may simply flow slowly in the due course of activity. Status information is not readily available and must usually be synthesized by contacting multiple resources. Often, one has to search and seek out where something is. As the process increases in complexity (number of products, number of languages), the project management becomes more difficult, time consuming, error prone, slow and expensive. At some point, automation becomes necessary.

Fortunately, some solutions are emerging.

The In-House System. Although not very interesting to the outsider, many Stage D organizations have developed their own systems to deal with automating the workflow. The benefit of an in-house solution is that it can be tailored to meet the exact needs of the organization and its process.

The solutions are server based or Web based. The project manager defines a project, describes the objects that are elements, assigns owners and describes the sequence of operations. The main objects that traverse the system are the translatable content files. If there is an in-house translation group, they are simply included in the process. However, many times an outside translation vendor is used, so that can create a “hole” in the process where it’s difficult to integrate the function directly.

The key benefit derived from these systems is visibility. One can simply inquire and receive a wealth of information about where any given object is, what’s happening to it, who is doing it and when completion is expected. Also, these systems can track cost information, which has historically been very difficult to get for translation/localization projects. So many different groups are involved (inside, outside, contractors) that consolidating financial data has proved to be elusive.

Vendor-based System. The translation service bureau, or vendor, has historically had problems that are very similar to the Stage D groups: many clients, many languages, many projects simultaneously. Projects involve using both in-house and contract resources. All projects are time-critical, and costs absolutely must be tracked. So, not surprisingly, many have developed very sophisticated systems to deal with managing the workflow. Their systems allow things like price estimates to be generated and sent to prospective customers, automated client-viewable status information, and complete start to final QA process management. Savings resulting from greater efficiency accrue both to the customer of such a vendor and to the vendor itself.

The customer merely drops sample files into an “inbox” along with a job description and gets a price quote and schedule for the work. After confirming and approving the quote, a complete set of translatables is placed in the inbox. Completed work (or interim versions for review and testing) is subsequently delivered to them in an “outbox.” The customer can log on at any time and see whatever status information he or she needs. The vendor runs a much more efficient business by having control over its resources and workflow. But the customer still has a problem of managing the rest of its process in-house. If the customer also has an in-house system, it would ideally be connected to the vendor’s system to simplify the overall process.

Hosted System. Slowly but surely, the combination of turn-around requirements, cost-control needs and overall complexity is drawing innovators to the table. New and demanding applications are emerging, which fuel the incentive to supply solutions. Picture a Web-based catalog sales organization with a global audience. Each item’s product description in the catalog must be presented in multiple languages; there are thousands of items; new items post daily; and the translations are needed very quickly — in hours instead of weeks.

What would be good is if the writer of the product description were able to drop it into an automated system that routed it to standby translators, who rapidly return translated copy, which the system posts on the Web site for viewing. Similarly, larger projects could flow automatically through a system where the customer and all involved vendors had some kind of pre-set financial understanding.

This is exactly what is emerging in the market today: systems hosted on the Internet, much like an ASP, where as an authorized user you can greatly simplify your process and project management. Although some evolution is sure to further improve the arrangements, you can deal now with systems that have automat ed the process of quotations, implementing translation resources, reporting, and delivery of final goods. There may be some limitations as to the selection of resources, language and so on, but the current offerings are significant.

Future System. Today, we create a product with a given language comprising its user interface or other textual content. The process whereby we convert our product’s content to be displayed in additional languages is becoming more sophisticated and simplified. Ultimately, we must evolve to a point where language becomes simply an attribute of the works that we create, as opposed to adding languages on as an afterthought.

We will see fully automated processes connected to the point of origin of the works we create. This way, you simply declare when you want a project to begin. When you type something in English, the system routes it through a process that returns it in each target language. When you have completed your work, it is already published also in each of your target languages. Sound incredible? We have most of what’s needed for this already, and projects are in the works to push it further.

In Summary

Demand for translation is likely to continue to increase. We can expect projects to be larger and larger, increasingly complex, and will be demanded in ever-shrinking timeframes. Technology is coming to our rescue, but only by carefully analyzing our process requirements can we implement meaningful solutions. Much of the challenge lies in managing the process and in minimizing pressure on critical path resources such as translators. We have not yet fully integrated all of the tools and technologies available into a seamless top-to-bottom solution, but we are on a realistic path to do so.
In this article I will discuss some of the most prominent features of translation memory (TM) tools. My intention is not to give you a list of “this tool or feature is better than that one.” What I want to show is how the different tools have realized the basics of a TM system.

At first glance, all the TM tools promise the same things: recycling of repetitive text parts, time and cost savings, better consistency, data exchange and so on. That sounds nice and easy. But as soon as you have decided that you need to work with one or more of these tools, you will have to spend some time evaluating and comparing features and processes.

Although I will not be able to save you all the time and effort of looking at the tools for yourself, this article might give you an idea what to look for when you do your evaluation. Please always keep in mind that there is no “one best tool” that can satisfy all the needs that a translator or project manager might have. One tool might serve your purpose better than another, depending on the processes you have, the file formats you work with or the user preferences you have.

I have looked at some of the TM tools on the market today: Déjà Vu by Atril; SDLX by SDL; Translator’s Workbench by TRADOS; Transit by STAR; TRANS Suite 2000 by Cypresoft; and Wordfast by Champollion Wordfast, Ltd. I have not included IBM Translation Manager, as this tool is no longer being marketed.

In this article, I will take a look at these features: TM model; translation environment; translation memory exchange format (TMX); statistics such as word count; fuzzy matching; and special elements such as abbreviations and acronyms.

### TM Models

Before going into individual features, I would like to introduce the two different models that are being used in TMs: the database model and the reference model.

The database model saves the source segment and its translation as one unit, a so-called translation unit, often abbreviated as TU. But it saves this unit out of context. The reference model does not save the segments and their translations in one place but references their position within the documents themselves. That is, if you want to see the segment in context, you can call up the reference material and take a look. But you also have to decide before you start a project which files or previous projects you want to use as reference material, whereas a database will just grow with every translation unit you save.

### Translation Environment

Let me start with a comparison of the most visible part of all translation tools — the translation environment. There are basically two different ways: working with an existing editor in addition to the TM tool and working within the TM tool.

Some tools (TRADOS Translator’s Workbench, Wordfast) use Word as an editor which simplifies the translation of all file formats that can be comfortably displayed in Word. But as soon as other file formats come into play, they will need to be prepared (converted) for use with Word or they have to be translated with a different editor altogether.

The other tools have their own editors. They import the files to translate and show them either in the form of a two-column table (one segment per table cell) or in different windows (one for the source language and one for the target language). They also differ in how they show formatting information and file structure information. Structural information is either omitted altogether or shown within protected tags. Formatting information can appear as tags or partly also in WYSIWYG mode.

### File Formats

Each tool provides a list of file formats that it supports. It is basically possible to translate any format with any tool. It really depends on the filters or conversion utilities that are offered by...
the manufacturer or the possibility to create or customize them.

A tool that uses its own editor, such as Transit, SDLX, Déjà Vu and TRANS Suite, will offer a list of supported file formats during the creation of a project. But some tools only allow one format per project; with others you can choose several formats to be included in one project.

A tool that uses Word as its editor, like Translator’s Workbench and Wordfast, doesn’t need any conversion or filtering as long as the files to be translated can be opened comfortably in Word. As soon as other file formats have to be processed, the tool either offers another editor, such as TagEditor from TRADOS, or the files need to be converted so that they can be opened in Word for translation. This also means that after translation the files have to be converted back again.

Although the display has not much to do with the actual translation features of a tool, this is one of the most important factors for many translators when they decide which tool to use. Another decisive factor is that many customers today know about translation tools or even use one of them in-house and want the translator to use that tool as well. But with the advent of TMX (translation memory exchange format), this can be circumvented to some extent.

TMX

TMX is a format that is used to exchange data among TM systems. So, if you want to reuse a TM that was produced with one tool in another one, TMX is the way to go. TMX is an XML-based description of the information that is contained in a TM system. There are three different levels of TMX compliance, which I will list here very briefly. (To see more information on the TMX specifications, go to the site www.lisa.org/tmx)

- TMX level 1 includes a description of text information only.
- TMX level 2 includes a description of text information and formatting information.
- TMX level 3 includes a description of text, formatting and tool-specific additional information. Some tools offer the possibility to create or customize them.

Examples of Six Translation Environments

- Déjà Vu
  - Text-only table
- Transit
  - Windows for source and target text, partly with WYSIWYG formatting
  - Copy of source text is overwritten during translation
- SDLX
  - Table, formatting partly WYSIWYG
- TRANS Suite 2000
  - Source and target window
- Windows for source and target text, partly with WYSIWYG formatting
- Copy of source text is overwritten during translation

Adams Globalization

For 20 years, Adams has helped Fortune 500 companies like Dell, Polycom and AMD to access the world. From Software Localization and Testing to Technical Translation, we can help you explore a world of possibilities.

For More Information:
Toll Free: 1.800.880.0667
Tel: 512.821.1818 • Fax: 512.821.1888
Email: sales@adamsglobalization.com
www.adamsglobalization.com
of adding information such as project name, file format, customer name, name of translator and so on. Most tools on the market today comply with level 2.

Even though the TMX standard exists, this still does not guarantee that all data can be brought safely from one tool to the other, as different tools use different structures for encoding RTF formatting information, subsegments such as footnotes or index entries. So, even though a tool can export text and formatting information into a TMX file, another tool might not be able to read it correctly, and you will end up with text only in your new memory.

**Examples**

The following sentence in an RTF file — This sentence contains different formatting information — will be represented differently in TMX depending on the tool used, as you can see in the accompanying code example.

Level 3 has not been implemented in any tool yet, as it demands the encoding of tool-specific information such as project name, customer name and other additional information with XML tags. But as different tools offer different customizability for such information, it is difficult to define how to encode it. For example, one tool offers a dialog where the user can define as many additional fields as he or she wants. Another tool offers only a limited number of fields. Here, the import from one into the other might result in loss of information.

Most tools also offer an import/export feature for the format of other tools' manufacturers. But as new tools pop up all the time and as manufacturers tend to work on their own formats — improving or migrating completely to another format — it is not possible for any tool to support all the other formats on the market. That was one of the reasons for the development of TMX.

**Statistics: Word Count**

How to calculate the word count (or line count respectively) may be one of the most disputed issues in the world of translation, since it is usually the basis for cost calculations. Here are some examples of how differently words are counted with different tools.

**Word.** Word counts everything between spaces as a word. A list of stand-alone symbols such as §, %, & , / creates a word count of four words. Numbers are also counted as words. Terms linked with a hyphen or a colon, such as pre-translation or human: machine, are counted as one word each.

**SDLX.** SDLX offers two different ways to count words: like Word or SDLX specific. The SDLX-specific counting includes numbers and special characters as words but not the symbols of bulleted or numbered lists, whereas Word counts these as words, too.

**Translator's Workbench.** This tool counts only translatable text. No stand-alone numbers, no symbols. A word here consists of at least one letter surrounded by spaces. Most TM tools count differently as they concentrate on translatable text. Stand-alone numbers or symbols are, therefore, usually not counted as words; or they count hyphenated terms as two words. In the

---

**TMX Representation Differs Between Tools**

**xTranslator's Workbench Example**

```xml
<tu>
  <tuv lang="EN-US">
    <seg>This <ut>{\b }sentence<ut>}</ut> contains <ut>{\i }different<ut>}</ut> <ut>{\ul }formatting<ut>}</ut> information.</seg>
  </tuv>
  <tuv lang="DE-DE">
    <seg>Dieser <ut>{\b }Satz<ut>}</ut> enthält <ut>{\i }verschiedene<ut>}</ut> <ut>{\ul }Formatierungen<ut>}</ut>.</seg>
  </tuv>
</tu>
```

**SDLX Example**

```xml
<tu xml:lang="EN-US">
  <seg>This <bpt i="1" x="1">1</bpt>&lt;1&gt;sentence&lt;/1&gt; contains &lt;bpt i="1" x="2">2</bpt>&lt;2&gt;different&lt;/2&gt; &lt;bpt i="1" x="3">3</bpt>&lt;3&gt;formatting information.&lt;/3&gt;&lt;/seg>
</tu>
<tu xml:lang="DE">
  <seg>Dieser &lt;bpt i="1" x="1">1</bpt>&lt;1&gt;Satz&lt;/1&gt; enthält &lt;bpt i="1" x="2">2</bpt>&lt;2&gt;verschiedene&lt;/2&gt; &lt;bpt i="1" x="3">3</bpt>&lt;3&gt;Formatierungen.&lt;/3&gt;&lt;/seg>
</tu>
```
Word Count Examples

<table>
<thead>
<tr>
<th>Test Segment/Tool</th>
<th>Word</th>
<th>Translator’s Workbench</th>
<th>Transit</th>
<th>SDLX</th>
<th>Dέjà Vu</th>
<th>TRANS Suite 2000</th>
<th>Wordfast</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. item</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2. point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• bullet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• list</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12356</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>The symbols are $, % and &amp;.</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Abbreviations like e.g. or i.e. in one segment.</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Hyphenated two-word expression.</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

“Word Count Examples” table, I will give you a few examples. This table illustrates what the word counts are for a few sample texts when using different tools.

As you can see, the statistics might look very different depending on the type of text that you are analyzing. And if you have sub-segments such as index entries or footnotes, the differences are even more visible.

Fuzzy Matching and Match Rates

Fuzzy matching is the technology that is used to find segments in the TM similar to those that have to be translated. The percentage of similarity between segments in the document and source segments in the TM defines the recyclability of the TM.

Most TM tools will create a statistic of match rates either with a separate analysis feature or automatically during import of the files to translate into the TM system. But because the fuzzy matching algorithms are a little different for every tool, the match rates in the analysis statistics are not really comparable.

Looking at the match rates for each tool as shown in the “Match Rate Examples” table, you will find that they are all set up differently.

Handling of Special Elements in a Segment

Abbreviations. Most tools recognize the common abbreviations of the language that is set as source language by internal lists. That means that even though they encounter a dot after a word, they do not end the segment there because they know that this is an abbreviation. Abbreviations that are not recognized can either be added in a dialog or with an external text file containing a list of unknown abbreviations.

Some tools also offer a list of words after import of the files to translate where the user can define which of those words should be handled as abbreviations.

Acronyms. Acronyms are words consisting only of uppercase letters, such as WYSIWYG. Some tools can recognize acronyms as units that do not need to be translated. Acronyms are even substituted automatically by some of the tools.

| Source segment 1: Use the acronym TM here. |
| Source segment 2: Use the acronym MT here. |

During translation of the second segment, some tools automatically substitute the acronym from the source segment for the one that was saved in the TM.

Summary

For this article I just did some small tests with simple RTF files, and there are many more things to consider when you are working with files coming from a DTP application or tagged files in HTML, SGML or XML. Take your time when you evaluate the tools and use some real-life files, not the sample files that are provided with the tools. They are great for getting to know how the tools work, but they do not give you the real-life picture.
Technology and the Freelance Translator

Jonathan T. Hine, Jr.

A Tale of Tom

Tom had just reached 55 km/hr down his favorite mountain road when he felt the pager buzz against his back. At the bottom of the hill, he turned his bicycle onto a side street and checked the number in the pager window. It was a translation company in Paris. He found a quiet area near a picnic table and called them back with his cellular telephone.

A half-hour later, he took a shower while his computer booted up. When he sat at his desk, he connected to the Internet and opened his e-mail. The message from Paris contained the specifications of the job and a two-megabyte compressed file attached to it. While the file downloaded, he opened his accounting software and started an estimate/job order form. Within a half-hour, he had opened the files, inspected them, verified the client’s word count and sent his estimate as an e-mail attachment to the client. He decided not to use dictation software for this assignment, so he put a CD of Schubert’s Third Symphony on the stereo and went to work.

The assignment consisted of engineering reports to the main office of a multinational oil company. Tom had done some of these the year before, so his translation tool had them in memory. He opened the tool, imported the job and examined it on the left side of the screen. He moved a few segments together and split out some others. Then he ordered the tool to pre-translate. While the tool worked, he opened the paper mail he had brought in from his bicycle ride. It took the tool five minutes to pre-translate all the reports, but it found matches for more than half of the segments in the document. It was time to edit the pre-translation.

Tom loaded a technical dictionary in the CD-ROM drive and launched his browser. Scrolling down, he compared each segment of source text to the target text and edited the target text as needed. Each time he finished a segment, a single keystroke ordered the tool to pre-translate all future segments using what Tom had chosen so far and to put his cursor in the next segment.

The next day Tom finished the document in the translation tool. He ran the code checker to make sure no formatting codes were lost and then the spell-checker. Finally, he exported the file back into Word and printed a copy for proofreading. The proof-reading went faster than it used to before he started using the translation tool, but he always managed to find something on paper that he had missed on the screen.

He entered the corrections in the document and attached the final version to an e-mail. His accounting software converted his estimate into an invoice and posted an entry to Accounts Receivable for him. He attached the invoice to an e-mail to his client’s accounting manager. While he was on-line, he checked his bank balance and saw that a wire transfer from another customer had come in. He downloaded that information into his accounting software and called it a good day’s work. He decided to ride to his favorite pizzeria to meet some friends and relax.

Same Job – Different World

Not much has changed about what translators like Tom (a fictional character) do since Jerome translated the Christian Bible into Latin. How we do it has changed a lot: for example, how we receive the source material and send the translation and our invoice to the clients (communications and administration); how we look up answers to our questions (research); and how we draft the translation and check our work (production and quality control).

Communications

Being there when the client calls is crucial. The translator who answers the phone or calls back first gets the assignment and keeps the client. How we achieve this depends on our individual lifestyles, our working environments and our budgets.

Whether the translator is working from home or an office, a voice telephone line has to stay clear. This may mean having a second telephone line for the computer and fax or a DSL (Data Subscriber Line) for the computer. My old-fashioned fax machine is reliable and works with the computer turned off. But today services such as Efax (www.efax.com) convert faxes into e-mail attachments. Incoming faxes are free; outgoing are inexpensive. Electronic faxes can be enlarged, rotated and copied to the clipboard. Electronic fax services in “wired” countries, such as www.freefax.it in Italy, allow the subscriber to give clients a local telephone number for faxes. Advertising-supported services such as www.freefax.com offer free outgoing faxes.

Cellular telephones have become standard issue. Some telephone companies offer a feature that allows the subscriber to forward the office phone to the cell number when running errands or working off-site. Some translators use the cell phone as their primary office phone.

Pagers come as separate units or are incorporated in cell phones. They let a customer leave a number or a message. The voice mail feature offered by some telephone companies will dial the pager when someone leaves a message. Pager signals can reach in areas where cell phone service fades.

Once the initial contact is made, often by telephone, most communications take place by e-mail. There are two elements in using e-mail and the Internet: the Internet connection and the various programs that reach out to the world through that connection.

The connection itself is the job of the Internet service provider (ISP). To check e-mail occasionally, a telephone line with a modem may suffice. Translators who rely on Internet sources for research and terminology checking (through Eurodicautom, yourDictionary.com, subject-area Web sites and so on) may need to be on-line for long periods. They may need a DSL or cable connection. Cable TV companies, telephone companies and cellular companies are competing vigorously for small office/home-office (SOHO) business. Translators and language service companies in populous metropolitan areas may need help choosing among the offers.

The programs running through the Internet connection include mainly the browser (Netscape, Internet Explorer and so on) and e-mail (usually with the browser, but also Eudora, Mulberry, Outlook and others). Then there are the applications that come with built-in Internet links, many of which support the business itself.
Administration

Taking care of business requires integrating the same technologies that translating does, plus on-line banking, e-commerce and support software. Armed with a credit card and a current browser, the wired freelancer sends invoices, pays bills and purchases what the business needs. Modern accounting software programs, such as Quicken, Money, Peachtree and others, include integrated browsing to various services over secure links. They also generate invoices in text or HTML that may be attached to e-mail.

At the very minimum, a spreadsheet program such as Excel and a presentation program such as PowerPoint should be in the freelancer’s kit. Often, translation assignments come as combinations of Word files with embedded spreadsheets as tables.

As we become increasingly dependent on electronic media, the stakes involved in a hard-drive crash, a fire or a power failure become even more important. It is time to include a CD read/write (R/W) drive, often called a “CD-burner,” in the computer suite. CD-burners offer a cost-effective way to back up files. They make it possible to store everything on the computer safely in a distant location.

An uninterruptible power supply (UPS) is also a good investment. An economical one can protect equipment from power fluctuations and provide time to save files and shut down gracefully during a blackout.

Finally, no discussion of technology can pass without mention of anti-virus software and intrusion protection. Translators and companies that use continuous connections like DSL or cable, in particular, need to install firewalls and keep their virus checkers updated. Symantec, makers of Norton AntiVirus, and Network Associates, which sells McAfee, are two of the best-known companies in the field. Those who do not know what a virus, a worm, a hoax and a firewall are or how to recognize a suspicious attachment should find out immediately.

Translation Tools

The first group includes products such as Déjà Vu and TransCorpora, which help the human being translate efficiently. They perform pre-translation and present the translator with a split screen: source text on one side, target on the other. Our mythical Tom was using such a tool. Localization project management software often includes translator modules that can be sent to the freelancers participating in the project. TRADOS, SDLX and STAR Transit are examples. These three also offer freelance editions which translators or freelance project managers can buy and use independently.

With translation tools, the source document must to be in an electronic format. This poses a problem for freelancers who perform a craft-like service on older paper documents, such as evidence in legal cases, or work with clients who do not provide electronic files. However, as attaching word-processing files to e-mail becomes more common, more freelancers will turn to translation tools to maximize efficiency.

Most translation tools suitable for freelancing or outsourcing projects work by accumulating strings (segments) of matched source and target texts into a translation memory (TM). The translator or the project team usually creates the TM by aligning segments of text in already-translated document pairs and saving the results. The tools also feature terminology databases. These are matched sets of source and target terms with additional information for sorting and filtering (parts of speech, gender, subject area and so on).

When ready to work on a new translation, the freelancer or project manager imports the source document into the tool. The latest tools will import from most major office application suites (Office, Lotus, Corel and so on). They embed codes in the imported document to preserve the formatting of the source document while working on the translated text.

The tool displays the imported file on half the screen, thus allowing the user to make any changes needed before pre-translation. For example, the tool may have started a new segment after each period in an abbreviation or numbered list. The user can join the parts of the list or the letters of the abbreviation, or change the segmentation rules and re-import the file.

When the source file looks good, the user orders the tool to pre-translate. The tool compares each segment with the segments in
When it finds a match, it puts the target segment for the match next to the source, filling up the other side of the screen segment by segment. Segments with no match may be left blank or may have the source text inserted automatically. Fuzzy matching allows the tool to retrieve partially matched segments if perfect matches are not found. For example, only the year or some numbers may be different in an otherwise perfect match.

Next, the translator edits what the tool has presented. This can be tedious or easy, depending on how many segments in the source document found a good match in the TM. The source and target segments are aligned on the screen, so comparing them and scrolling down are physically easy for the human being. As the translator scrolls to the next segment pair, the tool can automatically save the work, update the TM and pre-translate any segments still ahead. As the work proceeds, more segments from this particular document go into the TM. The tool presents more matched segments as the translator scrolls down.

The tool maintains its own file for all this work, so the source document is never modified. This allows a project manager to send the imported file and the TM and terminology database to a freelance translator who also has the tool. It can be sent at any point in the process, so who does the importing, checking and pre-translation can depend on the human being’s individual involvement in the project.

These tools work best when the source material is consistent in style and terminology and when there is enough of it to become repetitious. My own accuracy and speed increased just by having the split screen keep my work aligned in front of me and by having the tool pre-translate portions before I got to them. TM can also be shared, transmitted and updated by different people working on a common project. This facilitates consistency between translators.

For translators and companies that need to manage terminology, several programs are devoted to that task. They do not feature TM and pre-translation, but may provide cost-effective support to those working with paper source documents. LogTerm 3 is an example.

**Other Tools for Translators**

The other group of tools includes desktop publishing (DTP) programs and dictation software. DTP programs — Quark, Ventura, MS Publisher and others — perform the layout and typesetting needed for finished print documents such as this magazine. Technical communicators in many fields, including translators, may need these tools for particular clients. The DTP programs pick up where the word-processing programs leave off, so freelancers should know all the features of their own word processors — Word, WordPerfect, Lotus Notes, for example — before buying DTP software.

Dictation software replaces the old tape-recording that was transcribed by a typist in the old days. Before the personal computer, this was how high-volume, very specialized translators worked. At 125 words/minute with a familiar source text, a translator could dictate 40,000 words/day. Even today, it is a cost-effective way to produce the first draft of a translation. Dragon Naturally Speaking is probably the best-known of these programs.

Both groups of tools can also be used while working in the office programs with which they are compatible. For example, I can call up a translation tool database (terminology or TM) from MS Word, check a choice and insert it from the tool into the Word document directly. Similarly, I can have Dragon Naturally Speaking running in the background and dictate directly into the Word document. How well this integration works depends on the versions of each application and the amount of RAM, clock speed and free disk space available to the user.

**Changes in the Works**

Today, we see all this technology being combined and rearranged in many ways. For example, personal digital assistants (PDAs) are not only replacing bulky agenda books and telephone directories but are also integrating the cell phone and pager and providing portable PC technology. Notebooks and laptop computers allow freelancers to work anywhere, whether traveling for business or just working somewhere different. Wireless technology is becoming affordable and widespread, thereby alleviating the need to find a telephone connection for the computer on the road.

However, let us not forget the most amazing technological marvel of all.

**Timeless Technology**

What is at the center of this wired SOHO workplace? A few dozen kilometers of mostly water and organic material performing untold miracles of human thought. Some technology in the office may not be spectacular, but becomes vitally important to the long-term health of the human beings it serves. We should read about lighting, monitor height, wrist position and ergonomics and take a good look at where we work.

For example, is there adequate sunlight? If not, I would use full-spectrum light bulbs in some of the lamps. We will spend most of our working lives in the chair at the workstation, so we need the most comfortable, highest quality armchair we can afford. It took me two upgrades over four years to afford the chair I use now.

Tom probably set a good example for us by making sure his work did not keep him off his bicycle. We have to make the technology work for us and not the other way around. That way we will be there for ourselves, our clients and our families for many years to come.
Finding the Right Language Technology Tools

David Shadbolt

If your organization is new to language technology, this article introduces you to a few language software tools that help reduce the costs of globalization and improve delivery of different language versions to market. For the purposes of this article, the tools are categorized as machine translation (MT); internationalization; localization — embracing both translation memory (TM) and computer-aided translation (CAT); and, finally, multilingual content management, including multilingual workflow.

Before choosing tools for internal use or a localization vendor with those tools, Andrea Böhringer, marketing manager at STAR Group, suggests an organization needs to “look at the number of files that needs translating, as well as information within graphics for translating. It would need to consider timeframe, file management, publishing platforms and whether to use SGML or XML. I also recommend training technical writers in the techniques of writing and formatting to meet the requirements of TM processes, as well as building a terminology pool for source and target languages before starting with translation. If possible, implementing a workflow system allows automated handling of routine tasks.”

Machine Translation

After in-depth evaluation, an organization may decide that MT will provide sufficient quality to meet sales and customer service requirements, particularly if it has large volumes of content or if real-time translation is essential. In essence, MT scans a Web page that is in another language, parses the sentences into words and matches those words against custom dictionaries for comprehension before constructing translated sentences. Translation quality depends on the level of grammatical correctness and unambiguity of the source language coupled with MT technology that contains bilingual dictionaries containing all the source words. Although MT produces only gisting comprehension, the benefits of low cost and real-time translation have led to an enormous amount of documents translated with MT.

Founded in 1968, SYSTRAN claims that more than 400,000 Web sites translate over six million Web pages per day using its MT technologies. The company has recently undergone a complete redesign of its architecture to improve quality. Among many improvements, the redesign has modularized the code, converted to declarative programming, become fully XML compliant and made its dictionaries more accessible and intuitive. Reba Rosenbluth, director of corporate sales at SYSTRAN, says, “SYSTRAN offers three levels of translation quality. To begin with, there is gist quality, then first-level customization which is the ability to build your own dictionaries with Intuitive Coding Technology. The next level is publishing quality for near perfect translations, used for large-scale corporate customization projects. Often people do not know what they want, and we have to help them make the correct choice through a highly interactive process. It’s an education process, and the marketing incorporates the education. The bottom line is that we have to supply what the customer wants. We supply metrics to our customers, and if we say that we can provide 70% or 80% or 90% coverage of the language, we deliver.”

Another MT vendor is SDL International. Its Enterprise Translation Server (ETS) can process three million words per hour (50,000 per minute) and, claims Hedley Ree-Evans, marketing director at SDL, “has an application program interface that is much easier to implement than our main competitors, as well as a user-friendly interface for dictionary maintenance, which plays a crucial role in improving real-time translation quality. MT deployment is mainly as a gisting agent; where no translation was conducted before. For example, ETS works extremely well with the Lotus Sametime collaborative computing application (for secure, internally hosted chat and file sharing) in real time. The other obvious application is in processing fast, short shelf-life material, such as e-mail, where it is possible to read gisted translations, for example, to establish the importance of customer or partner mailing through a customer support center. It also plays an increasing role in the pre-translation productivity area, where MT dictionaries can be upgraded from terminology within TM databases and vice versa. We have clients deploying news-based processing in this way. This approach is also beginning to save SDL substantial translation time and costs.”

“The bottom line,” Rosenbluth says, “is that the demand for multilingual content is growing. Translators can’t keep up with the demand, and organizations cannot always afford the time and money to obtain high-quality translations, especially for large projects. If you take a million pages at $40 per page for translation, even a customized MT solution pays for itself very quickly.”

If MT is considered too costly, there are other alternatives. Without needing specialized language fonts or complex translation software, Langoo.com is providing a free service to over 70,000 registered users, 60% of them outside the United States, who are conducting between 25,000 and 30,000 transactions daily in 95 languages. Users can use any PC or Web device to compose and send e-mail messages, chat, send and receive electronic greeting cards and create personal Web pages in their own languages. Gary Kral, executive vice president at Langoo.com, says, “Our plan was to prove the technology, create a critical mass and then license the software component that makes up our system for commercial use. We are already doing that. For example, VeriSign.com is using our software to enable users to search and register multilingual domain names.”

For Web sites, if the international audience has a reasonable grasp of English, there is always automatic content enrichment (ACE) by RichLink, which associates words and phrases with dynamically created pop-up portals. Marc Bookman, CEO at Sentius Corporation, makers of RichLink, believes that for markets where English is a second language, ACE provides a useful option, particularly if “companies translate their top-level messages and use it for larger volumes of content.”

Internationalization

If an organization decides to deliver truly locale-specific products or applications, then it should consider internationalization. Just as templates and style sheets make it easier to...
produce each issue of this magazine, internationalizing the code base of a software product will expedite localization. Fully internationalized code facilitates easier handling of all types of character sets and display formats, such as currencies, dates and language-specific programming data for localized versions. The process usually begins with an architectural and systems analysis at the source-code level. Many language vendors, such as SDL, Rubric and Weblocalize, will conduct an evaluation and prepare a cost analysis if requested; however, they tend to partner with companies that specialize in this. A few of these companies have tools for code identification, retrieval, conversion and compliance with Unicode. For example, LingoPort has its Global Investigator that provides collaborative and batch environments for searching source code for internationalization errors, including embedded strings, image references, locale-limiting programming methods and/or functions, and customizable programming patterns. Basis Technology has its Rosette Globalization Platform that contains tools such as Language Identifier and Language Analyzer. One of the most interesting components is the Rosette Core Library for Unicode, which enables software engineers to quickly add support for over 150 of the world’s languages to their applications.

Steve Cohen, executive vice president at Basis Technology, points out that “one of the main categories for internationalization is resolving errors arising from the processing of information that is either input by the user, such as a search term or mailing address, or obtained from some other source, like a file or database. These processing steps in a program often make assumptions about the language, locale or nature of the text. In order to properly localize an application, the source code needs to be re-engineered so that it can handle text in another language. The Rosette Globalization Platform is primarily used for the processing and analysis of text, but we also offer engineering, source code audit and international quality assurance services that cover the other aspects of internationalization.”

Amazon.com has used Basis Technology products and services to internationalize its core platform and to launch its Japan Web site. Cohen says, “Amazon.com estimates that it will now be able to roll out new language versions in one third of the time it would have taken had it not internationalized. This is a significant improvement in Amazon’s marketing efforts, especially when put alongside its recent announcement that it expects Japan to be its biggest market outside of the United States.”

Localization

Creating the illusion that a product is a home-based brand is the purpose of localization. This includes translating, integrating culturally appropriate graphics and accommodating local currency and means of purchase, although other factors will need consideration. Specialized localization tools, such as PASSOLO 3.5 from PASS Engineering, Germany, and CATALYST 4.0 by Alchemy Software Development of Ireland, have proven their worth. Projects handled by PASSOLO can contain any number of programs and dynamic linked libraries (DLL) that, in turn, can be translated into many languages. The software can simultaneously use several general or project-oriented glossaries that the user can create individually. It is also possible to look up translations in a current translation list or in translation lists of other programs. The search considers similar texts as well as exact matches. “These so-called fuzzy matches increase the efficiency of the translator,” says Florian Sachse, managing director of PASS Engineering. “Our software also contains a dialog editor, which allows the translator to see his or her work in the correct context. Dialog layouts can be adapted to the translated passages without any danger of accidentally deleting or changing existing elements or structures. All of this improves translator productivity, especially when a simulator allows project managers and software developers to check the application for any errors in hard-coded text, insufficiently sized text buffers, and similar programming errors.”

Alchemy’s CEO Tony O’Dowd says, “Localization tools, such as our CATALYST 4.0, save time and money by seamlessly integrating editors and tools for translators, localization engineers, QA specialists, project managers and software developers. Our tool has many features that shorten production and ensure that translations quickly with no need to correct the alignment. Then they can query them when they wish, just as they would query a terminology database. The bitexts are created in HTML format, they can be easily displayed on the Internet. It’s also an inexpensive tool compared to TM and other similar tools.”

Most software localization tools use the reference model, which is the one where a source segment and its translation are not saved in one place. It is referenced within the document itself. TM, on the other hand, uses the database model which saves both the source segment and its translation as one unit. (See “Comparing Basic Features of TM Tools” in this supplement.) TM databases thus grow with every translation saved. Some TM tools are TRADOS’ Translator’s Workbench, STAR’s Transit, Atril’s Dējā Vu and SDL’s SDLX. It is worth noting that localization tools typically provide the option of having an interface to a TM system. For example, PASSOLO provides interface options for TRADOS and STAR.

With translated sentences accumulating in the TM database, all repetitions are adopted automatically in subsequent projects. The translator thus only works through new text segments. TRADOS TM technologies, which receive extensive QA texts across 72 languages, have achieved the greatest market penetration over the past two decades. David Lawson, chief technology officer at TRADOS, says, “Over the years, many billions of words have been processed by our technology, and our customers have helped us perfect our algorithms for all the strange exceptional cases found in human languages.”
STAR’s Böhringer points out the benefits of Transit XV: “Our Transit XV supports a total of 160 languages and language versions. In comparison to conventional translating, you can save up to 80%. TM works best when combined with other terminology management tools. We have TermStar to ensure uniformity of terminology, particularly in the case of product information, operating manuals, software descriptions and other documentation. TermStar XV considerably improves this structured work. Formulations that have been recorded once and selected do not have to be entered again. During the process of translation, Transit XV continuously searches for terminology that already exists in the database. Expressions found are color-highlighted in the text, displayed in the terminology window as ‘defined’ and can be taken over in the text by means of a simple keystroke combination.”

Some vendors prefer to use the term computer-aided translation tool (CAT) rather than TM. Atril calls its Déjà Vu “an integrated computer-assisted translation system,” and SDL calls its SDLX 4.0 a CAT tool. Rees-Evans says, “We call it CAT because we believe the terminology is more informative to those who are not professionals in the industry. SDLX is, in fact, a TM system in the true sense and has some major features, such as permitting the same segmentation rules for all file formats, which is important for leveraging between different file types. In addition, through the use of the OpenTag format, cross leveraging of different file types is much easier than with competitor products because within OpenTag all formatting information is the same. Also, with SDLX, we can apply one or more TM databases to any number of files.”

SDL believes that SDLX 4.0 CAT is a complete productivity solution, consisting of a suite of tools that includes the SDL Project Wizard for guiding the user through the file preparation process, conversion and, after translation, the subsequent rebuilding back into the original file structure. SDL Edit enables translators easy access for tasks such as identifying in TM full and partial matches, terminology checks and documentation preview; SDL Align allows for the updating of a TM or creation of a new one, matches existing translations and validates alignment using built-in heuristics, as well as a terminology database and management tools.

While every vendor tends to believe the features of its TM (or CAT) bring something new to the table, Daniel Gervais, executive director at MultiCorpora R&D Inc., believes that traditional TM and other CAT tools have failed to deliver significant benefits, except for technical documentation with a high degree of repetition. “Their limitations,” Gervais says, “stem from their dependence on whole sentence repetition, their loss of translation context and the prohibitive labor-intensity of building the initial TM database. Our MultiTrans is a corpus-based translation support system that has demonstrated superior productivity for all types of content, including non-technical administrative correspondence and marketing content where whole sentence repetition is almost non-existent.”

“MultiTrans is based on a full-text, indexed and aligned collection of previously translated content (the corpus),” Gervais continues. “A corpus of a million words can be built in less than an hour and can be ready for immediate use by translators and terminologists. Unlike TM databases that require tedious manual effort to ensure perfect alignment, the full context approach of MultiTrans allows users to easily identify at a glance and correct alignment as they work. MultiTrans allows users to rapidly access words, terms, expressions or sentences in any language contained in multiple corpora. Expressions and their previous translations are always provided in the context of their original documents, thereby providing valuable usage and style guidance.”

“MultiTrans also deploys a sophisticated search and matching algorithm,” he says, “to automatically compare an entire new source document to multiple open corpora (and terminology repositories) and to
Language service bureaus generally have a small team of in-house translators and an extensive pool of freelancers. Internal quality control and project management processes are not always so effective when bringing freelance translators into a project, which was the motivating factor in the development of Televend’s T-Remote Memory according to the company’s CEO Jean-Didier Boucau. "For the big projects with several translators for each language," Boucau explains, "we were always confronted with a duality: Do we have to work with external translators and lose time in project management, repetitions and leveraging and so on, or do we have to bring the translators on site and pay for travel and accommodation? As there wasn’t a tool that met all our requirements, we developed our own. With T-Remote Memory, a company can harness the power of the Internet to enable several translators, based anywhere in the world, to work on-line as a team. They can run several translation memories at the same time, whatever their format with no restrictions. With full and easy integration of translators into projects, the virtual office becomes a reality."

Estimating the cost of a project is an important factor in deciding whether or not much localization is required. There are tools for facilitating this task, including WebBudget 3 XT which is an innovative addition to them. The software produces reliable content analysis in practically any language simply by selecting the corresponding character set in the "profile" located under the options tab.

Choosing a Vendor
Choosing a vendor will depend on the company, the product and/or application in need of localization and the size, number of languages, formats and other considerations such as IT infrastructure and confidentiality aspects. Televend’s Boucau says, "I should, of course, first check if the vendor has the linguistic and technical expertise. If this is established, I think the most important thing to do is to try to work with people with whom there is a real fit. The human aspects remain the most important constraint or advantage during a project. It is very important in the beginning to define good and clear communication lines so that the process can be optimized and respected. It should always start with a ‘consulting’ phase in both directions, and this should lead to agreed-upon process and planning."

Böhinger at STAR believes an organization should "search for a translation service agency which is able to translate documents with the same process and tools it intends using in the future."

Future of Language Tools
Language software tools will continue to evolve. Boucau sees the sector developing in two key ways: "The tools will most probably be more and more Web-related. Translation remains something quite expensive and still — in most industries — non-strategic. A big part of the costs is still related to project management, communication and engineering. The power of the Internet will certainly solve a major part of those aspects. Another key area is TM systems. I think they will become commodity products and the major interest of this technology will be to couple them with MT. Linguistic engineering should improve a lot of applications such as alignment, terminology extraction and so on."

Vendors are already moving in that direction — SDL with its SDLWebFlow enabling users to choose TM, MT or other options via portal and TRADOS with the architecture of its new Language Management Platform based around the Internet, relational database and XML technologies. Lawson says, "I think you’ll find that we’re really listening to our customers, so you’ll see us produce the technology solutions that the market requires." But while existing language tools can improve productivity and speed time-to-market for products, the usual caveat applies if an organization intends on purchasing or licensing for its own internal use. Is the anticipated ROI realistic for the organization?

We surveyed more than 200 companies and collected information about their language tools. The next two pages show the respondents and their products. For more information, including product pricing and features, go to www.multilingual.com/ldetail

---

Multilingual Content Management

A few major localization vendors have developed multilingual workflow systems (MWS) to manage language processes from start to finish. An MWS should contain a global management system (GMS) for managing multilingual content, along with TM and MT. Content management systems (CMS) from vendors such as Stellent, Vignette, Documentum, J.D. Edwards and FatWire were not originally designed to synchronize translation and localization of content. Most have not developed content integrators to work in partnership with GMS from vendors such as GlobalSight and TRADOS (formerly Uniscape). SDL also has its own software product for this application. Rees-Evans says, "In the case of our own GMS, SDLWebFlow, we have a purpose-built content integration application, SDLNexus, that is customized to each client's installation."

A capable GMS requires components that will automatically detect changes in databases and content repositories; extract content; separate translatable content from any format code; leverage earlier translations from TM databases; notify translators and administrators; and manage the project according to pre-defined workflow (business rules) procedures.

In addition to a GMS, "achieving the seamless deployment of both reusable high quality and fast ‘gisted’ translations, thus achieving the maximum value and return on investment, a multi-vendor workflow is essential," contends Rees-Evans. "To obtain 50% to 60% savings of time in the hand-off of target languages, there needs to be a simple yet flexible method to set up and manage workflows — for all content types, languages and resources. Finally, to accommodate ad hoc translation work that is not being extracted automatically, an MWS needs a front-end or ‘portal’ application that will have an easily installable or browser-based application for the user to define the work requiring translation and to ensure that it is routed to the RT server or TM workflows. This facilitates easy translation of discretionary jobs like Word files, but is a fairly simple form of workflow."

---

LANGUAGETECHNOLOGY
Language Technology Products and Developers

This list is divided into the three major groups that form the language-technology bridge. Integrated multilingual workflow systems and programs are listed in the blue section, translation tools and suites of applications in red, and engineering tools and computer resources in purple. Some of the suites and larger systems may include individual applications or elements that belong to another category. We strongly recommend viewing the fully detailed list on our Web site (www.multilingual.com/ltDetail) for a better understanding of the basic functions of these products as well as their use and relationship to other products and to multilingual technology as a whole.—Jerry Luther

<table>
<thead>
<tr>
<th>Company</th>
<th>Web Address</th>
<th>Product/Suite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Technology Products and Developers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated Workflow Systems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALCHEMY USA Software House</td>
<td><a href="http://www.alchemysoftware.ie">www.alchemysoftware.ie</a></td>
<td>ALCHEMY CATALYST 4.0</td>
</tr>
<tr>
<td>AppTek, Inc.</td>
<td><a href="http://www.apptek.com">www.apptek.com</a></td>
<td>QuantumCMS</td>
</tr>
<tr>
<td>Asia Communications</td>
<td><a href="http://www.apptek.com">www.apptek.com</a></td>
<td>WebBudget</td>
</tr>
<tr>
<td>ATRIL Software SL</td>
<td><a href="http://www.atril.com">www.atril.com</a></td>
<td>GlobalSight System4</td>
</tr>
<tr>
<td>Basic Language Systems Corporation</td>
<td><a href="http://www.multilingo.com">www.multilingo.com</a></td>
<td>Multilizer 5.1</td>
</tr>
<tr>
<td>BridgeTerm, a division of InfoGraffti Inc.</td>
<td><a href="http://www.bridgeterm.com">www.bridgeterm.com</a></td>
<td>Obtree C4 Content Management</td>
</tr>
<tr>
<td>CIMOS</td>
<td><a href="http://www.cimos.com">www.cimos.com</a></td>
<td>InsightBuilder</td>
</tr>
<tr>
<td>Cyepressoft</td>
<td><a href="http://www.cypressoft.com">www.cypressoft.com</a></td>
<td>Enterprise Translation Server, SDL Localization Suite, SDLWebFlow</td>
</tr>
<tr>
<td>Ectaco, Inc.</td>
<td><a href="http://www.ectaco.com">www.ectaco.com</a></td>
<td>TermStar XV, Transit XV Professional, TWS, WebTerm 5, STAR Proactive</td>
</tr>
<tr>
<td>EWGate Pte. Ltd.</td>
<td><a href="http://www.ewgate.com">www.ewgate.com</a></td>
<td>T-Remote Memory</td>
</tr>
<tr>
<td>GalTech Soft</td>
<td><a href="http://www.galtechssoft.com">www.galtechssoft.com</a></td>
<td>LTC Organiser</td>
</tr>
<tr>
<td>I.R.I.S., Inc.</td>
<td><a href="http://www.irissusa.com">www.irissusa.com</a></td>
<td>MultiTerm ix, Corporate Translation Solution, TRADOS 5.5 Freelance, TRADOS 5.5 Service Provider Suite, Uniscape GXT</td>
</tr>
<tr>
<td>IBM Corporation</td>
<td><a href="http://www.ibm.com">www.ibm.com</a></td>
<td>UniSite Content Manager</td>
</tr>
<tr>
<td>Institute for Language &amp; Speech Processing</td>
<td><a href="http://www.ilsp.gr">www.ilsp.gr</a></td>
<td>Web500 CMS Professional</td>
</tr>
<tr>
<td>International Systems Consultancy Ltd</td>
<td><a href="http://www.isc.com.au">www.isc.com.au</a></td>
<td></td>
</tr>
<tr>
<td>ITR International Translation Resources Ltd.</td>
<td><a href="http://www.itrblackjack.com">www.itrblackjack.com</a></td>
<td></td>
</tr>
<tr>
<td>J.D. Edwards</td>
<td><a href="http://www.jdedwards.com">www.jdedwards.com</a></td>
<td></td>
</tr>
<tr>
<td>Julia Emily Software</td>
<td><a href="http://www.lexicool.com">www.lexicool.com</a></td>
<td></td>
</tr>
</tbody>
</table>

Translation Tools and Suites

<table>
<thead>
<tr>
<th>Company</th>
<th>Web Address</th>
<th>Product/Suite</th>
</tr>
</thead>
<tbody>
<tr>
<td>AppTek, Inc.</td>
<td><a href="http://www.apptek.com">www.apptek.com</a></td>
<td>LocalSphere, TextFinder, TranSphere</td>
</tr>
<tr>
<td>Asia Communications</td>
<td><a href="http://www.apptek.com">www.apptek.com</a></td>
<td>KEY, MacKey, KEY CJK Server, KEY CJK Web Server, KEYTIP</td>
</tr>
<tr>
<td>ATRIL Software SL</td>
<td><a href="http://www.atril.com">www.atril.com</a></td>
<td>Deja Vu</td>
</tr>
<tr>
<td>Basic Language Systems Corporation</td>
<td><a href="http://www.multilingo.com">www.multilingo.com</a></td>
<td>MultiLingo</td>
</tr>
<tr>
<td>BridgeTerm, a division of InfoGraffti Inc.</td>
<td><a href="http://www.bridgeterm.com">www.bridgeterm.com</a></td>
<td>ProMemory</td>
</tr>
<tr>
<td>CIMOS</td>
<td><a href="http://www.cimos.com">www.cimos.com</a></td>
<td>An-Nakel Translator</td>
</tr>
<tr>
<td>Cyepressoft</td>
<td><a href="http://www.cypressoft.com">www.cypressoft.com</a></td>
<td>TRANS Suite 2000</td>
</tr>
<tr>
<td>Ectaco, Inc.</td>
<td><a href="http://www.ectaco.com">www.ectaco.com</a></td>
<td>Partner Translator, Universal Translator for Pocket PC</td>
</tr>
<tr>
<td>EWGate Pte. Ltd.</td>
<td><a href="http://www.ewgate.com">www.ewgate.com</a></td>
<td>EWT transate</td>
</tr>
<tr>
<td>GalTech Soft</td>
<td><a href="http://www.galtechssoft.com">www.galtechssoft.com</a></td>
<td>WordPoint, Genesis 11</td>
</tr>
<tr>
<td>IBM Corporation</td>
<td><a href="http://www.ibm.com">www.ibm.com</a></td>
<td>IBM WebSphere Translation Server</td>
</tr>
<tr>
<td>Institute for Language &amp; Speech Processing</td>
<td><a href="http://www.ilsp.gr">www.ilsp.gr</a></td>
<td>TrAID Align, TrAID Term Memory, TrAID Translation Memory DB</td>
</tr>
<tr>
<td>ITR International Translation Resources Ltd.</td>
<td><a href="http://www.itrblackjack.com">www.itrblackjack.com</a></td>
<td>ITR Blackjack</td>
</tr>
<tr>
<td>J.D. Edwards</td>
<td><a href="http://www.jdedwards.com">www.jdedwards.com</a></td>
<td>J.D. Edwards Content Manager</td>
</tr>
<tr>
<td>Julia Emily Software</td>
<td><a href="http://www.lexicool.com">www.lexicool.com</a></td>
<td>lexicool.com, Lingo - Translator’s Assistant</td>
</tr>
</tbody>
</table>
## Translation Tools and Suites (continued)

<table>
<thead>
<tr>
<th>Company</th>
<th>Web Address</th>
<th>Product/Suite</th>
</tr>
</thead>
<tbody>
<tr>
<td>linguatec language technologies GmbH</td>
<td><a href="http://www.linguatec.de">www.linguatec.de</a></td>
<td><strong>Translation Tools and Services</strong> — Home, Network, Office, Office Plus; e-Translation Server</td>
</tr>
<tr>
<td>LogOMedia Corporation</td>
<td><a href="http://www.logomedia.net">www.logomedia.net</a></td>
<td><strong>LogoMedia Translate</strong></td>
</tr>
<tr>
<td>MetaTexis Software and Services</td>
<td><a href="http://www.metatexis.com">www.metatexis.com</a></td>
<td><strong>MetaTexis</strong></td>
</tr>
<tr>
<td>MultiCorpora R&amp;D Inc.</td>
<td><a href="http://www.multicorpora.com">www.multicorpora.com</a></td>
<td><strong>MultiTrans</strong></td>
</tr>
<tr>
<td>Practiline Software</td>
<td><a href="http://www.practiline.com">www.practiline.com</a></td>
<td><strong>PractiCount &amp; Invoice</strong></td>
</tr>
<tr>
<td>PROMT</td>
<td><a href="http://www.e-promt.com">www.e-promt.com</a></td>
<td><strong>PROMT</strong></td>
</tr>
<tr>
<td>RHS Group, LLC</td>
<td><a href="http://www.translate.com">www.translate.com</a></td>
<td><strong>RWS Localization Tools, JCAT</strong></td>
</tr>
<tr>
<td>Sentius Corporation</td>
<td><a href="http://www.sentius.com">www.sentius.com</a></td>
<td><strong>RichLink</strong></td>
</tr>
<tr>
<td>Synthema</td>
<td><a href="http://www.synthesa.it">www.synthesa.it</a></td>
<td><strong>Terminology Wizard</strong></td>
</tr>
<tr>
<td>SYSTRAN</td>
<td><a href="http://www.systransoft.com">www.systransoft.com</a></td>
<td><strong>SYSTRAN Family of Desktop Products, SYSTRAN Corporate Solutions</strong></td>
</tr>
<tr>
<td>Terminotix Inc.</td>
<td><a href="http://www.terminotix.com">www.terminotix.com</a></td>
<td><strong>LogiTerm</strong></td>
</tr>
<tr>
<td>TM SYSTEMS</td>
<td><a href="http://www.tm-systems.com">www.tm-systems.com</a></td>
<td><strong>Translation</strong></td>
</tr>
<tr>
<td>Translation Experts Ltd.</td>
<td><a href="http://www.tranexp.com">www.tranexp.com</a></td>
<td><strong>InterTran, MobileTran, NeuroTran, PalmTran, PocketTran, WordTran</strong></td>
</tr>
<tr>
<td>TranslationWave.com</td>
<td><a href="http://www.translationwave.com">www.translationwave.com</a></td>
<td><strong>TranslationWave</strong></td>
</tr>
<tr>
<td>TwinBridge Software Corporation</td>
<td><a href="http://www.twinbridge.com">www.twinbridge.com</a></td>
<td><strong>Chinese Partner Professional, Chinese Partner, Japanese Partner, Korean Partner</strong></td>
</tr>
<tr>
<td>Ultragluinga</td>
<td><a href="http://www.ultragluinga.com">www.ultragluinga.com</a></td>
<td><strong>Ultragluinga Dictionaries</strong></td>
</tr>
</tbody>
</table>

## Engineering Tools and Computer Resources

<table>
<thead>
<tr>
<th>Company</th>
<th>Web Address</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agfa Monotype Corporation</td>
<td><a href="http://www.agfamonotype.com">www.agfamonotype.com</a></td>
<td><strong>WorldType Multilingual Font Library</strong></td>
</tr>
<tr>
<td>AIT AG</td>
<td><a href="http://www.visioc.com">www.visioc.com</a></td>
<td><strong>Visual Localize</strong></td>
</tr>
<tr>
<td>Atia Ltd.</td>
<td><a href="http://www.atia.com/languagestudio">www.atia.com/languagestudio</a></td>
<td><strong>Language Studio</strong></td>
</tr>
<tr>
<td>Babylon Ltd.</td>
<td><a href="http://www.babylon.com">www.babylon.com</a></td>
<td><strong>Babylon-Pro, Babylon-Corporate</strong></td>
</tr>
<tr>
<td>Bantam Software Limited</td>
<td><a href="http://www.bantam.co.nz">www.bantam.co.nz</a></td>
<td><strong>TransLocale</strong></td>
</tr>
<tr>
<td>Champollion Wordfast Ltd.</td>
<td><a href="http://www.champollion.net">www.champollion.net</a></td>
<td><strong>Wordfast</strong></td>
</tr>
<tr>
<td>Connexor Oy</td>
<td><a href="http://www.connexor.com">www.connexor.com</a></td>
<td><strong>Connector Machinese OEM</strong></td>
</tr>
<tr>
<td>Convera</td>
<td><a href="http://www.convera.com">www.convera.com</a></td>
<td><strong>RetrievalWare</strong></td>
</tr>
<tr>
<td>em2 solutions</td>
<td><a href="http://www.em2-solutions.com">www.em2-solutions.com</a></td>
<td><strong>Font Augmented Internet Resource System (FAIRY)</strong></td>
</tr>
<tr>
<td>Everlasting Systems Limited</td>
<td><a href="http://www.everlastingsystems.com">www.everlastingsystems.com</a></td>
<td><strong>SGIL for VC, SGIL Software I18N Tools</strong></td>
</tr>
<tr>
<td>Fontlab Ltd.</td>
<td><a href="http://www.fontlab.com">www.fontlab.com</a></td>
<td><strong>AsiaFont Studio, FontLab, ScanFont</strong></td>
</tr>
<tr>
<td>Helicon Software Development</td>
<td><a href="http://www.helicon.co.at">www.helicon.co.at</a></td>
<td><strong>Helicon Translator</strong></td>
</tr>
<tr>
<td>IDC Global (Europe) Limited</td>
<td><a href="http://www.idcglobal.com">www.idcglobal.com</a></td>
<td><strong>dataFloat</strong></td>
</tr>
<tr>
<td>Langoo.com</td>
<td><a href="http://www.langoo.com">www.langoo.com</a></td>
<td><strong>Lan go Input, Display and Storage application</strong></td>
</tr>
<tr>
<td>LingoPort, Inc.</td>
<td><a href="http://www.lingoport.com">www.lingoport.com</a></td>
<td><strong>Global Investigator</strong></td>
</tr>
<tr>
<td>MARGI Systems</td>
<td><a href="http://www.margi.com">www.margi.com</a></td>
<td><strong>Presenter-to-Go</strong></td>
</tr>
<tr>
<td>Netrics, Inc.</td>
<td><a href="http://www.netrics.com">www.netrics.com</a></td>
<td><strong>Netrics Search</strong></td>
</tr>
<tr>
<td>Nstein Technologies Inc.</td>
<td><a href="http://www.nstein.com">www.nstein.com</a></td>
<td><strong>Nserver E-Publisher Suite</strong></td>
</tr>
<tr>
<td>ParaType Inc.</td>
<td><a href="http://www.paratype.com">www.paratype.com</a></td>
<td><strong>Multilingual Fonts, ParaWin</strong></td>
</tr>
<tr>
<td>PASS Engineering GmbH</td>
<td><a href="http://www.passolo.com">www.passolo.com</a></td>
<td><strong>PASSOLO</strong></td>
</tr>
<tr>
<td>Polderland Language &amp; Speech Technology</td>
<td><a href="http://www.polderland.nl">www.polderland.nl</a></td>
<td><strong>Flexicon, LIT, SmartMatch, spelling checkers</strong></td>
</tr>
<tr>
<td>Schaudin.com</td>
<td><a href="http://www.schaudin.com">www.schaudin.com</a></td>
<td><strong>RC-WinTrans</strong></td>
</tr>
<tr>
<td>Segue Software</td>
<td><a href="http://www.segue.com">www.segue.com</a></td>
<td><strong>SilkTest International</strong></td>
</tr>
<tr>
<td>WhippleWare</td>
<td><a href="http://www.whippleware.com">www.whippleware.com</a></td>
<td><strong>VB Language Manager</strong></td>
</tr>
<tr>
<td>WizArt Software</td>
<td><a href="http://www.wizart.com">www.wizart.com</a></td>
<td><strong>WizTom</strong></td>
</tr>
<tr>
<td>Zi Corporation</td>
<td><a href="http://www.zicorp.com">www.zicorp.com</a></td>
<td><strong>eZiTap, eZiText</strong></td>
</tr>
</tbody>
</table>

For complete information about these companies and technologies, see [www.multilingual.com/ltDetail](http://www.multilingual.com/ltDetail)
This supplement introduces you to the magazine *MultiLingual Computing & Technology*. Published nine times a year, filled with news, technical developments and language information, it is widely recognized as a useful and informative publication for people who are interested in the role of language, technology and translation in our twenty-first-century world.

**Language Technology**

From multiple keyboard layouts and input methods to Unicode-enabled operating systems, language-specific encodings, systems that recognize your handwriting or your speech in any language — language technology is changing day by day. And this technology is also changing the way in which people communicate on a personal level; changing the requirements for international software; and changing how business is done all over the world.

*MultiLingual Computing & Technology* is your source for the best information and insight into these developments and how they will affect you and your business.

**Global Web**

Every Web site is a global Web site, and even a site designed for one country may require several languages to be effective. Experienced Web professionals explain how to create a site that works for users everywhere, how to attract those users to your site and how to keep it current. Whether you use the Internet and World Wide Web for e-mail, for purchasing services, for promoting your business or for conducting fully international e-commerce, you’ll benefit from the information and ideas in each issue of *MultiLingual Computing & Technology*.

**Translation**

How are translation tools changing the art and science of communicating ideas and information between speakers of different languages? Translators are vital to the development of international and localized software. Those who specialize in technical documents, such as manuals for computer hardware and software, industrial equipment and medical products, use sophisticated tools along with professional expertise to translate complex text clearly and precisely. Translators and people who use translation services track new developments through articles and news items in *MultiLingual Computing & Technology*.

**Managing Content**

How do you track all the words and the changes that occur in a multilingual Web site? How do you know who’s doing what and where? How do you respond to customers and vendors in a prompt manner and in their own languages? The growing and changing field of content management and global management systems (CMS and GMS), customer relations management (CRM) and other management disciplines is increasingly important as systems become more complex. Leaders in the development of these systems explain how they work and how they work together.

**Internationalization**

Making software ready for the international market requires more than just a good idea. How does an international developer prepare a product for multiple locales? Will the pictures and colors you select for a user interface in France be suitable for users in Brazil? Elements such as date and currency formats sound like simple components, but developers who ignore the many international variants find that their products may be unusable. You’ll find sound ideas and practical help in every issue.

**Localization**

How can you make your product look and feel as if it were built in another country for users of that language and culture? How do you choose a localization service vendor? Developers and localizers offer their ideas and relate their experiences with practical advice that will save you time and money in your localization projects.

**And There’s Much More**

Authors with in-depth knowledge summarize changes in the language industry and explain its financial side, describe the challenges of computing in various languages, explain and update encoding schemes and evaluate software and systems. Other articles focus on particular countries or regions; translation and localization training programs; the uses of language technology in specific industries — a wide array of current topics from the world of multilingual computing.

Nine times a year, readers of *MultiLingual Computing & Technology* explore language technology and its applications, project management, basic elements and advanced ideas with the people and companies who are building the future.

To subscribe, use our secure on-line form at [www.multilingual.com/subscribe](http://www.multilingual.com/subscribe)

Be sure to enter this on-line registration code: sup51
In addition to the world’s leading Translation Memory system, TRADOS 5.5 provides you with professional language technology that extends your capabilities and enables your business to become more productive and profitable. Plus you have the advantage of working seamlessly with thousands of translators and customers around the world who already use TRADOS today.

Find out more today at trados.com or call 1-888-2-TRADOS or 1-703-683-6900 ext. 2640. Write to us at tsd@trados.com.