TRANSLATION • LOCALIZATION • INTERPRETING • GLOBAL COMMUNICATION

Language International

The business resource for a multilingual age

October 2001 www.language-international.com Vol. 13 No. 5

Gauls, Britons, and Pilgrims

Vivendi Universal's Anne Grill on localizing educational products. page 12

A Leisure Industry but a Serious Business

Why videogames don't always translate well into other cultures... page 16

The State of Flash

A multimedia expert on the localization of Macromedia Flash projects. page 30

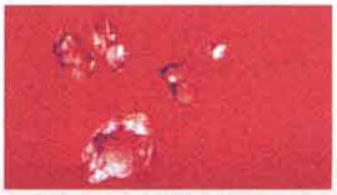
Translation Memory for Free

An interview with Yves Champollion, the creator of WordFast. page 36

Language International

The business resource for a multilingual age

Gauls, Britons, and Pilgrims:
Vivendi Universal's Anne Grill explains how educational software packages are adapted and localized.—**Page 12.**



State of Flash:
The tricks and tips of localizing Flash presentations used on the Web.—**Page 30.**

It's Free, But Does It Work?
A Bob Clark interview with the creator of WordFast, a free translation memory tool. —**Page 36.**

Editorial

News and Views

The Business of Language

Education Spotlight

"Hi, my name is Raphael" is the meaning of the Japanese title of this issue of Language International. In this issue, we focus on the translation, adaptation, and localization of multimedia products and games.

Opinion

Tools and Technology

State of the Art

Subscription Information
Subscriptions rates for 6 issues, postage included (airmail). VAT is required within the EU, excluding those with VAT-ID numbers.

Individual:* US$76.00, UK£ 42.49 (incl.VAT), NLG 144.16 (incl.VAT), NLG 136.00 (excl.VAT), EUR 65.42 (incl.VAT), EUR 61.71 (excl.VAT).

Institutional: US$110.00, NLG 220.00 (excl.VAT), EUR 99.83 (excl.VAT).

*Private subscriptions are for personal use only and should be prepaid. £ Prices valid only for UK residents; $ prices valid only for US residents. Residents of other countries should pay in Dutch Guilders (NLG) or Euros (EUR). Please send your order to:

John Benjamins Publishing Co.
P.O. Box 36224 •
1020 ME Amsterdam •
The Netherlands

ISSN 0923-182 X

From the Editors

Bert Esselink

Bob Clark

Not many people will dispute the fact that Language International is a rather diverse magazine. Often there seems to be no focus, no clear target audience, no specific regional attention, and it's definitely no a-z read for everyone. For example, in our June issue this year, you could read about GlobalSight's Web globalization management technology system, followed by an article on the threatening extinction of the Osage language. In August topics ranged from Internet B2B development in Europe to the translation work done at the European Commission. In this month's issue you will find similar extreme opposites...

The only thing the articles in Language International have in common is a distant or close relationship to *language*. But is this a bad thing? We at Language International don't think so. The "language industry" is one of the most diverse and culturally rich environments to be working in, and there are many things to learn that may not directly relate to your own area of expertise. For me personally a recent visit to South Africa was a huge eye-opener. Learning about the difficulties of revitalizing and preserving thirteen official languages puts a whole new perspective on the problems that localizers typically face, such as how to quickly resize dialog boxes. And hearing terms like Web globalization and double-byte enablement being interpreted into Sesotho is quite an experience in itself.

This issue of Language International is focusing on multilingual multimedia production. One of the feature articles is titled "A Leisure Industry but a Serious Business", which definitely applies to the localization and translation of educational and entertainment software products. Multimedia products have always been hard to localize. Not only are the levels of adaptation often much higher than the average technical manual or Web site, the software used to create or localize multimedia products is not always "localization-friendly".

In our Business of Language section we have an exclusive interview with Anne Grills, who is responsible for localization at Vivendi Universal, a leading global media and communications company. She speaks about the complexities of localizing educational products. In the same section, Flavia Grant from Eidos Interactive in the United Kingdom discusses the legal issues and content adaptations in the localization of computer games.

A more technical approach is taken in the articles on Macromedia Flash localization and Localization in the Wireless World. Flash is a technology used in many of today's Web-based multimedia products and animated Web pages, and can be hard to localize if not developed correctly from the beginning.

Finally, in the new Opinion section of Language International, you will find Janet Fraser's reply to Anthony Pym's article on academics versus translators from the August issue. Janet Fraser is Senior Lecturer in Translation at the University of Westminster and calls for academics with the appropriate skills and expertise to become involved in the organizations representing translators. Her contribution ends with a phrase that can also be applied to the readers of Language International: "Be bold and come on in: the water's lovely, and who knows, everyone might enjoy drying off together in the sun after a shared swim."

What's In It For Me?

The information published in *Language International* is based on the following regular "departments" and feature columns:

1. News & Views—A selection of important and interesting news items related to languages in the broader sense as well as reports on recent developments in the language and localization industry.
2. Business of Language—Case studies and in-depth discussions of the impact that globalization issues and language complexities have on international businesses.
3. Education Spotlight—Case studies and articles on universities and translation schools training tommorrow's language professionals.
4. Tools & Technology—Product reviews and comparative analyses of globalization technologies along with stories on the impact of new technologies on the language industry.
5. State of the Art—Various stories dealing with language and translation standards, research projects, and best practices.

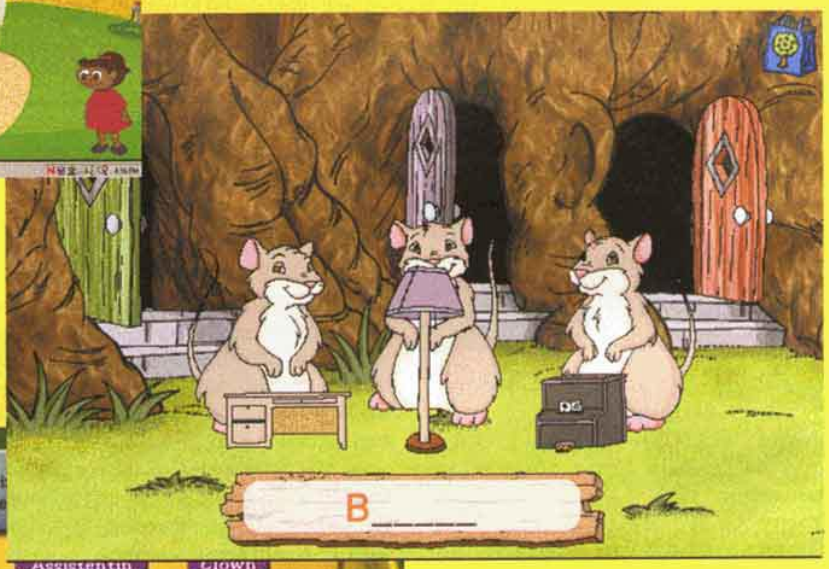
We are confident that this mix of news, bussiness issues, technology delopments, and in-depth articles on education and the art of working in the language industry will provide you with all the tools necessary to keep you "at the cutting edge."

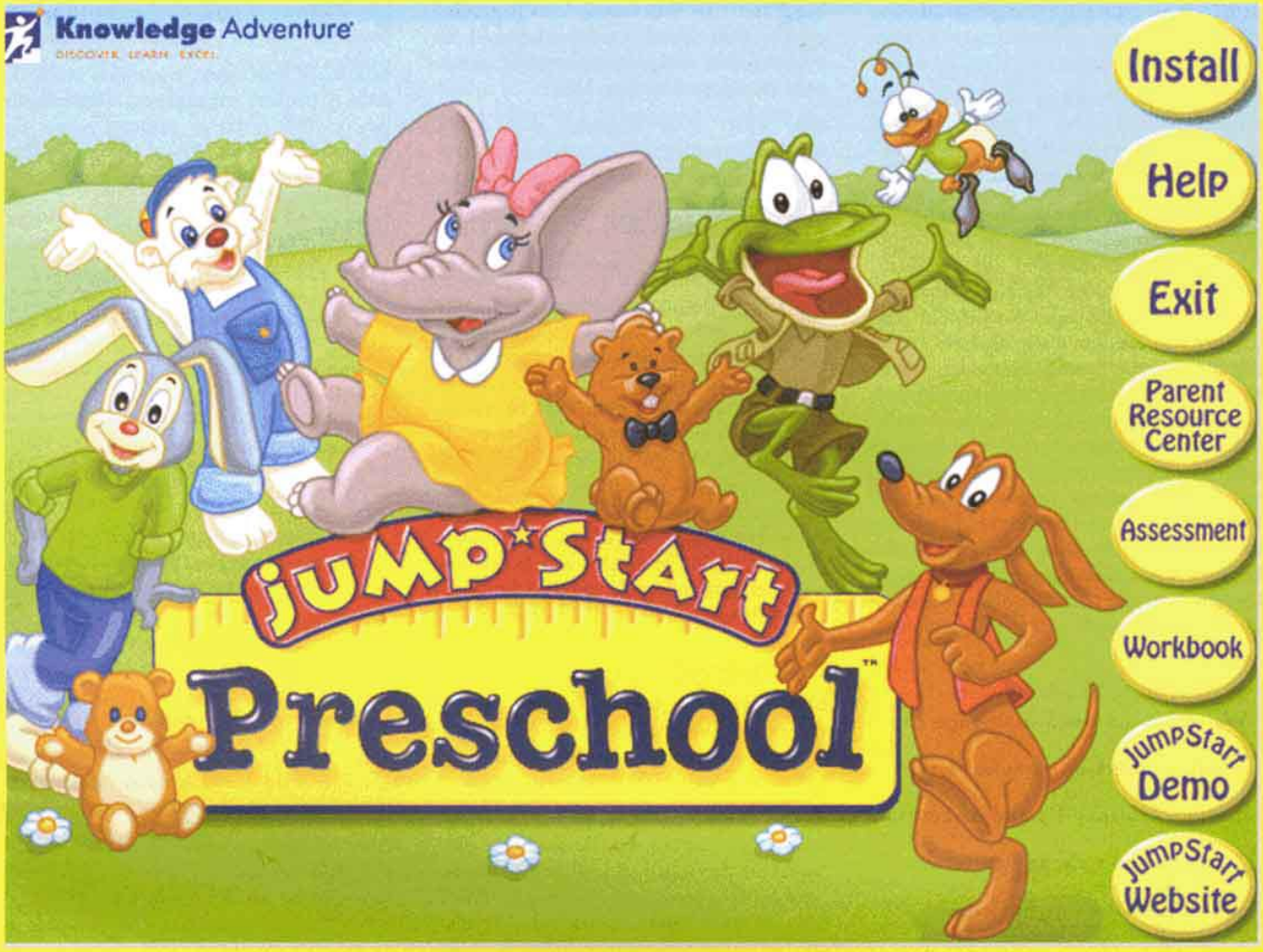
Your feedback and input continue to be critical to our success. **You** make our magazine. So, if you have a story that applies to any of the above categories, don't be shy. Send it in!

The Gauls, the Britons, and the Pilgrims

Vivendi Universal's Anne Grill on the challenges of localizing educational software packages.

by **Bert Esselink**





Vivendi Universal is a global media and communications company operating in a wide range of businesses, including music, TV and film, publishing, telecoms, and the Internet. Anne Grills is Content Manager for Vivendi Universal Interactive Publishing, the company formerly known as Havas Interactive. Anne is responsible for the localization of educational software packages, which is managed out of Vivendi's localization center in Dublin.

Language International met up with Anne at the Localisation Summer School in Limerick, Ireland, where she presented on "Managing Cultural Diversity in education and elearning."

Language International: Please describe how you got involved in the localization of multimedia products?

Anne Grill: I was a teacher for many years and was always very interested in the use of technology for educational purposes. I had also studied languages and cultural studies and so when I saw an ad for CUC, as we were then known, it seemed a natural step. I took a career break from my teaching job to move into the area of educational multimedia. I've ended up staying!

LI: What does your typical day look like?

AG: There really isn't a typical day I'm afraid. A lot of my daily tasks will relate to the kinds of titles we are working on. Generally though the first thing I do in the morning is check any mail that may have come in over night from the US or Canada.

The percentage that changes really depends on the suitability of the material to the target culture and curricula.

Then I will usually have a few meetings scheduled to discuss upcoming projects and projects actually in progress. I work with Content Department members throughout the day to assess localization issues and find possible solutions. There is a lot of consultation with other departments and management of resources. With 70 products coming through the Content department at the moment it's definitely busy!

LI: What is the localization center in Dublin responsible for?

AG: The localization center here in Dublin is responsible for localizing Vivendi Universal games, educational and edutainment titles. We organize the adaptation of content, translations, recording of audio, graphic rework, engineering, DTP, QA and project management of course; all aspects of localization in fact.

LI: What types of people are involved in the localization process?

AG: In Dublin we have all kinds of people. We have people with very specialized talents such as our Graphic Artists and Engineers. We have people from teaching backgrounds for our educational products and serious game lovers for testing titles, as well as the departments who work directly on products. Also, we have a great support team in our Finance and HR depart-

ments. It is an extremely dynamic atmosphere with diverse nationalities and great characters.

LI: What percentage of an educational software package is translated versus adapted or recreated for the local market?

AG: Actually we translate everything and then we begin the process of adaptation. Absolutely every element in an educational title is reviewed to ensure that it is appropriate for the target market. The percentage that changes really depends on the suitability of the material to the target culture and curricula. In a reading product, for example, we will change a lot of material due to the way phonics are taught in different countries and obviously anything with an historical element will be rewritten. More and more we are getting involved with our developers in the United States and France. The Content Department in particular gets involved at the design stage of a product to identify any issues that may be controversial or unsuitable for markets other than the original one. We have a strong focus in Dublin on internationalization.

We may even need to change the whole context of modules if they are very US specific and of course we need to change cultural references such as food items or historical and geographical references.

LI: Who decides what needs to be adapted for each individual target market?

AG: We send all titles to be reviewed by specialists in the target country. They review them firstly from a top-level perspective indicating the main areas for consideration. In conjunction with our out-of-house specialists the Content Department will then make recommendations to the project managers. Once the project team has been informed of the changes to be made and has assessed the implications, the subject specialists will look at each individual audio file and graphic to make changes and ensure that everything is adapted.

LI: What are the most common elements that need to be changed in educational and entertainment products developed in the US?

AG: Apart from language issues, (UK versus US spelling, for example, is very important when it comes to educational titles), the way numbers are written in mathematics varies from the US to European countries. The number 7, for example, is crossed in France but not in the US. Addition is horizontal in some countries and vertical in others. Images such as postboxes, police uniforms, cars on the right hand side of the road all need to be changed. We may even need to change the whole context of modules if they are very US specific and of course we need to change cultural references such as food items or historical and geographical references. For example, in a French developed product, there recently was a history module that referred to the Gauls in France. This was modified to teach UK children about the Britons and American children about the Pilgrims.

LI: Do you use translators for the localization work, or subject matter experts (such as teachers for the educational products)?

AG: We use translators purely to translate scripts, text graphics and the like. Some of them are specialized in translating for a younger audience, which helps greatly, and multimedia translators are used for gaming titles. We also, however, use subject matter experts for all our products. When we have cooking products, for example, we employ chefs to review and adapt the material, we have used genealogists for our genealogy titles and gardeners for our landscaping products.

LI: How do you deal with the various educational standards?

AG: Our in-house Content Specialists are very familiar with the educational curricula in the target countries and are used to the adaptations necessary. Of course educational standards vary and what is acceptable for one age group in France may be too easy or too difficult in Germany or the UK. We have a Content Database that we constantly update and that has material covering curricular and cultural issues and illustrates the different approaches to education around the world. It helps us to ensure consistency and keep up to date with all issues. The database is available to our developers too.

LI: Please describe the localization process for a typical multimedia product.

AG: A project manager is assigned and he or she initiates contact with the original producer. We receive a localization kit and all departments carry out a localization analysis. Budgets are finalized on the basis of the analysis and resources are assigned. Components are then translated, linguistically reviewed and content revised. Recording takes place in the target country and then all components are integrated and we begin the build and QA cycle.

Of course educational standards vary and what is acceptable for one age group in France may be too easy or too difficult in Germany or the UK.

LI: How do you localize the audio used in the products?

AG: Usually we translate the audio first and then we send the files to in-country experts who may be teachers or other subject experts. We train them on how to adapt material, supply them with the definitive list of changes we have agreed should be made and advise them on style and content. They then work through the audio files adapting or rewriting where necessary. It is very important that any changes to the files are cross-referenced with text graphics and Help files for example. The files go through a final check in Dublin before integration.

LI: What is the hardest product type to localize, and why?

AG: Educational products are probably the hardest to localize, and from a content perspective reading products are particularly difficult. France does not teach the alphabet the same way as the UK. Once material has been translated, for any language, rhymes will no longer agree, groups of words that began with the same letter will have changed, and songs and rhymes may no longer sound rhythmic. Of course sim-shipping our game titles across multiple platforms (PC, PlayStation, Game Boy), requires very close working relationships with the developers, intensive

periods of work and a lot of coordination to achieve.

LI: What defines a truly well localized multimedia product?

AG: If a product looks as though it has been developed for the target market then it has been successfully localized. Nothing should appear awkward and the user should not be distracted by something that seems out of place. He or she should not be aware of the fact that the title has been developed initially for a different country. If certain food items are familiar to children in France then they should appear in the French version. If a gardening product is for Germany then the soil samples mentioned should be appropriate to that country.

LI: What can be done by product developers to make localization easier?

AG: Product developers should aim to make their products as localizable as possible. Of course the product has to be appropriate for the 'mother' country too and I would not advocate an 'international' product but rather an 'internationalization friendly' product. The application needs to be coded in such a way that it adapts dynamically to the user's language and cultural conventions. Shared files, concatenation, and hard coded strings should all be avoided and all components should be resourced, held in external data files, to ease their subsequent localization.

Once material has been translated, for any language, rhymes will no longer agree, groups of words that began with the same letter will have changed, and songs and rhymes may no longer sound rhythmic.

LI: What advice can you give to other companies localizing multimedia products?

AG: They should try to have as much contact with the developers as possible and have input to the products at an early stage. There is nothing worse than being surprised with key elements that need to change but which are hard coded. Attention to detail is also something I recommend, particularly when it comes to children's product. It is also vitally important to have a good knowledge management system in place so that information and skills acquired in one localization process can benefit other projects.

LI: What references can you recommend to people who want to learn about multimedia localization?

AG: I enjoy reading Multilingual, Localisation Ireland, and of course Language International. I often consult the official educational websites of our target countries and www.lisa.org. For educational software localization in particular there isn't a lot out there and that is why it is so important that we constantly update our own process documents in-house and create and maintain documentation and resources for our developers and internal teams.

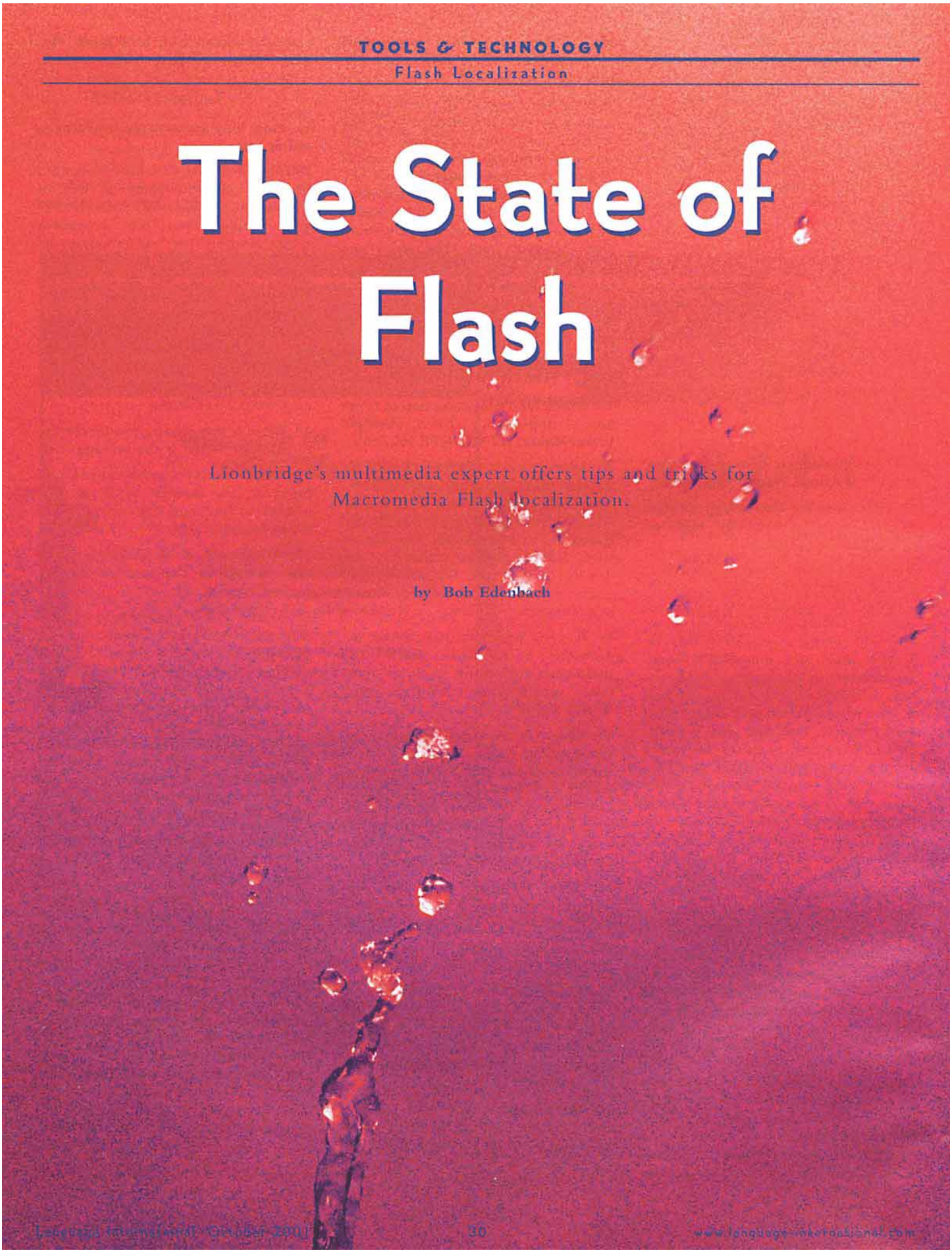
For more information on Vivendi Universal Publishing, visit www.vivendiuniversalpublishing.com.

Photo: Courtesy Vivendi

The State of Flash

Lionbridge's multimedia expert offers tips and tricks for Macromedia Flash localization.

by Bob Edenbach



Whether you love it or hate it, there is no arguing that the Flash technology has created quite a splash over the past two years. Its popularity surged with the release of Flash 4 in 1999, which came with a recognizable interface and an approachable proprietary scripting language called ActionScript. Web designers were quick to jump aboard and ride the wave of a streaming media revolution armed with an easy to use and comprehensive multimedia tool optimized for the Web. Within months, companies competed to have the slickest splash page, and words were literally flying off home pages as Web users flocked to download the Flash browser plugin to experience the latest multimedia experience.

There was a glut of extraneous Flash use and Web critics found themselves lurching for the "skip intro" button as fast as they fled from the infamous <BLINK> tag. Despite the outcries of HTML purists, Flash was a resounding success and, a year later, with the release of Flash 5 with enhanced scripting support and database compatibility, Flash had entrenched itself soundly within the Web culture. So what exactly is Flash, and, if it has become a cornerstone of the Internet, how do we deal with it in the face of the multilingual Web?

As a result, developers can create low-bandwidth multimedia presentations, Web sites, and applications.

Welcome to the World of Macromedia's Flash

Flash is a vector-based animation technology, generating a formidable coupling of high-impact streaming media and small file size. As a result, developers can create low-bandwidth multimedia presentations, Web sites, and applications. Flash files can be viewed independently as projector files or through a browser. Macromedia states that over 96% of Internet users now have some version of the Flash plug-in installed in their browsers. Unlike DHMTL and other scripting languages, Flash is cross-platform and cross-browser compatible, making it an attractive solution, skirting the tedium of writing browser-specific code for basic design and functional elements.

If you plan to localize Flash content into Asian languages, there are definite limitations and various guidelines you will have to follow.

There are many issues to localizing Flash, but the good news is that many of the techniques used to optimize the localization process will also reap benefits in the development stages. Essential questions for each Flash developer prior to localization include:

- Are you developing a Web site entirely in Flash or just complimented by Flash?
- Into which languages do you plan on localizing your content?
- Will you be creating a single multilingual file?
- Do you want the user to be able to view text in multiple languages, regardless of their operating system and language preference?
- Will your Flash files contain dynamic content from an external source?

If you plan on targeting European languages only, then Flash is a sensible alternative, as many potential localization issues are similar to those encountered when designing for the Web or a graphical user interface. A developer must take into account spatial issues due to text expansion, the application of text effects, the use of screenshots, currency and date formats, and choosing imagery that is culturally appropriate for the desired audience.

If you plan to localize Flash content into Asian languages, there are definite limitations and various guidelines you will have to follow. It is important to note that Japanese is the only double-byte language that is currently fully supported by Flash. As a result, dynamically generated text and text strings embedded within ActionScript will only be viable for Japanese and users will only be able to see this content on a Japanese system. This definitely should not inhibit you from creating Flash movies in Chinese or Korean, it just means that all your text will have to be "broken apart" (i.e. rendered or flattened).

Figure 1: This navigation bar is not optimized for localization, as the text abuts the sides of the buttons leaving no room for text expansion. As a result, the point size of the text will have to be considerably reduced, the buttons widened or the translation abbreviated.

Localization Process for Flash

The first stage is analyzing the Flash source (FLA) files. Obviously, all the textual elements that will need to be translated must be obtained in order to localize a Flash file. The organization and development style of these files will have a great impact on the time required for localization. In some cases, it may take hours to reorganize the symbol libraries and comb through the animation timeline to identify all the source text that needs to be localized. This includes text strings from static text fields within the FLA, as well as strings in imported graphics, text within action script, dynamic text imported from an external source and any text present in the audio

The key to smooth localization revolves around a systematic approach to both development and organization of the symbol library.

if there is narration. The use of the Movie Explorer will help you find hidden text, audio and graphics that appear in the timeline. Once all the text is extracted, it is then placed into a spreadsheet, translation table or text file, and placed into a translation memory tool for translation.

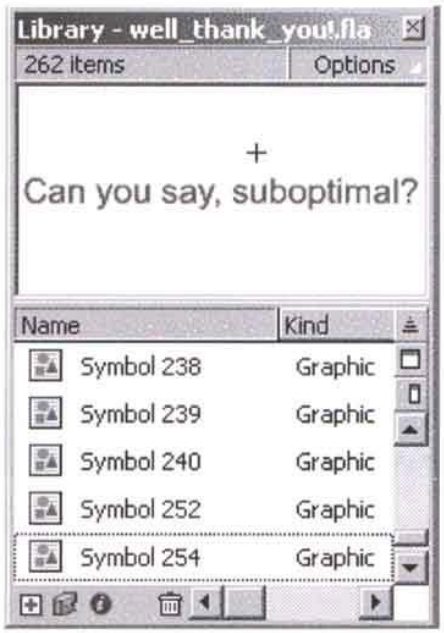


Figure 2: An unorganized library with unlabelled symbols.

After translation is completed, strings are exported from the translation memory tool for integration into Flash. If there is audio, it is recommended to send the translated script to the client for review and approval before recording, due to the cost and inefficiency of re-recording linguistic edits. After approval of the script and selection of a voice talent, the audio will then be recorded and edited. Simultaneously, the translated strings are inserted into the FLA file, and resized and aligned as necessary. The localized audio tracks are imported and synched to the animation within the movie or animation. The movie is then tested repeatedly to ensure the synching has been executed correctly and the files stream and render as smoothly as the original. The exported movie (SWF) is then reintegrated into the Web files or sent out as a projector file (EXE) for linguistic testing. Any edits are then made and the files are rebuilt and sent to the client for review.

Developing Flash for Localization

The key to smooth localization revolves around a systematic approach to both development and organization of the symbol library. Start by writing comments explaining actions within the timeline and functions within the ActionScript, and properly labeling layers, frames and symbols. Symbols are an incredibly powerful resource and should be used extensively. Basically, once a symbol has been created it can be used multiple times within a movie without adding to the size of the file. There are three different types of symbols: graphic, button, and movie clip. First create all textual elements as graphic symbols and align each symbol consistently to the stage.

The organization and development style of these files will have a great impact on the time required for localization.

Do not use live text in a movie that is not a symbol. In large movies it will be hard to find, especially if it is within a movie clip or triggered by a mouse action such as a rollover/hit state of a button. In addition, try to avoid using text in ActionScript. Often it is just as easy to develop the same functionality using graphic symbols and frame labels with ActionScript.

Once you have created all the text symbols and labeled them accordingly, place them in a separate folder from your graphics. Now you can begin creating your button and movie clip symbols, by placing the graphic symbols with text into your new symbol (this is called "nesting"). To create new states of your button in different frames for rollovers and hits, you can copy your first frame (right click) to the new frame and change the color of your graphic symbol by manipulating its tint or alpha value to the desired color. As a result, all text in the movie, including text within moving clips and buttons, will be updated and properly aligned when you localize the text in the graphic symbol. In addition, your Flash files are now easily editable for content updates and their libraries can be shared with other Flash movies.

Following a few simple guidelines will help you circumvent many potential localization catastrophes.

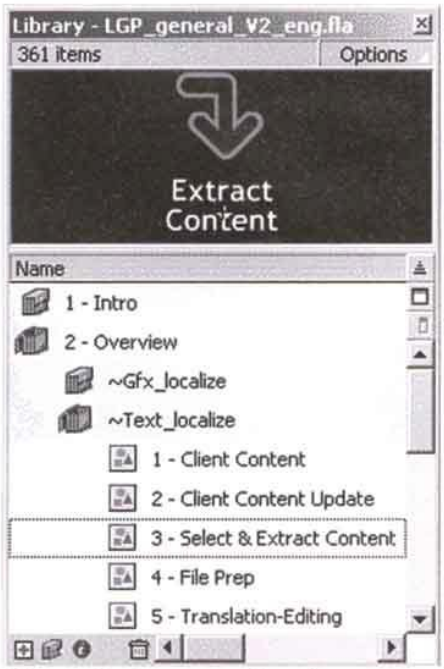


Figure 3: An example of an organized library identifying symbols with localizable text.

Refrain from using text effects dependent upon the use of single letter symbols combined to form text strings. When the movie is localized the translation for this text string will certainly contain a different number of letters, requiring redevelopment of the text effect or ruining the integrity of the effect entirely (for instance, the Korean translation may be only 4 characters long).

Asian text must be input on localized systems into your Flash file and converted to a symbol. It is important to create two versions of these symbols, one of which is editable and one that is "broken apart." There can be issues with some Asian font outlines (especially Traditional and Simplified Chinese) collapsing when you break apart the text, resulting in a filled object, so you must test and choose fonts carefully.

Following a few simple guidelines will help you circumvent many potential localization catastrophes. It will also allow

you to harness the full power of this application and the potential innovation in your design. The influence of Flash and Flash design with its compartmentalized views and alpha channels has gone far beyond the streaming media world, and entered standard forms of media as well. Its influence can be seen in television commercials, music videos, magazine ads, and billboards.

Bob Edenbach is currently the Manager of Web & Multimedia Services at Lionbridge Technologies in Framingham, MA. Prior to his three years at Lionbridge he was a graphic designer in Newport, RI. He can be reached at bob_edenbach@lionbridge.com.

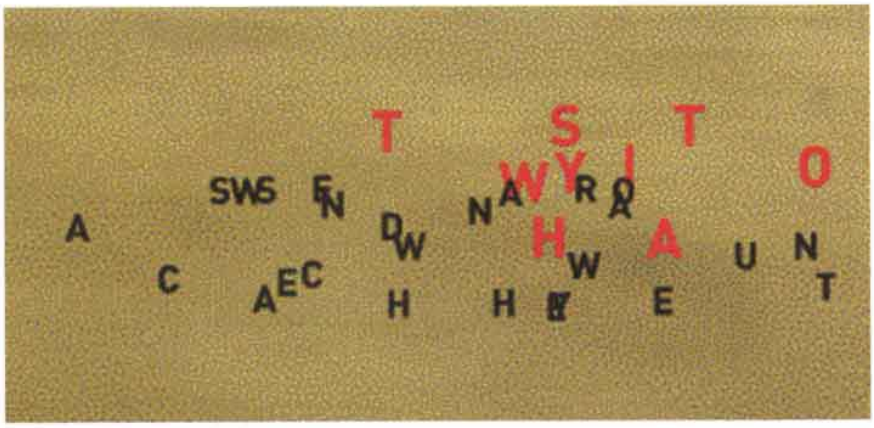
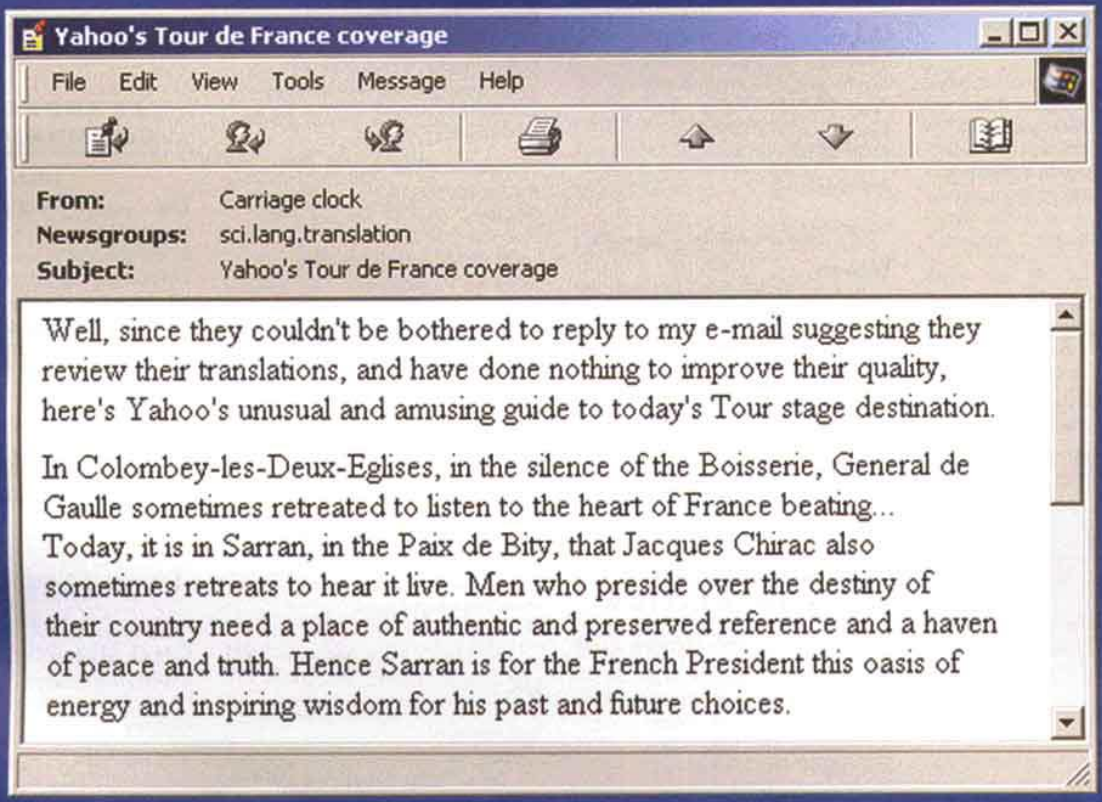


Figure 4: This popular text effect, where text blows apart into individual characters will require hours of rework for each language.

Bob Edenbach

Found on the Web:

Yahoo's Coverage of the Tour de France



WORDFAST®

It's Free, But Does It Work?

An exclusive interview with Yves Champollion, developer of a free translation memory tool.

by Bob Clark

The immortal words of John McEnroe, "You can't be serious!" pretty much sum up the general reaction to the initial appearance of WordFast on the email circuit several months ago. The skepticism was understandable. Here we have a full-blown translation memory system that appears out of nowhere, seems to be fully contained in Microsoft Word and it's free. Bert and I discussed including an item in *Language International* at the time and he, quite rightly, pointed out that over the years he had seen loads of macro-based "translation memory" solutions and none of them worked very well. I had a quick peek at the software and it seemed good enough to contact the developer, Yves Champollion, to find out more about it.

Champollion is not the sort of person you would normally associate with the development of translation software. Neither software engineer nor computational linguist, Champollion spent his early years drifting around the world, doing odd jobs and living by his wits. His Japanese-born wife was the one who "saved him from himself" eighteen years ago and dragged him back to France where he ran an import/export business dealing in Japanese fine art. And how did he get involved in translation? "I suppose it was inevitable," says Champollion. "An uncle four generations back, Jean François Champollion, was the one who cracked the code of the Rosetta Stone. I guess people figured that if he could do that, then I could probably translate automotive manuals. Translation is a tough job. You do the same thing over and over for years. It's not my nature at all.

Take any programmer who is in his mid-forties and you will find that he is an ex-New Age guy, an ex-Flower Power kid who sold his VW van and bought an Apple II in 1981.

That's why I went into developing a tool that helps you to translate. Translation memory is only one stage. My interest is actually machine translation."

So, how *did* this ex-drifter, with no conventional higher education manage to produce WordFast? "I am self-taught. In the 1980s I was a fan of micro-computing and I studied all sorts of programming languages on my own. Take any programmer who is in his mid-forties and you will find that he is an ex-New Age guy, an ex-Flower Power kid who sold his VW van and bought an Apple II in 1981. All these guys are reconstructed hippies. That's sort of what happened to me. In the early eighties I sold my sleeping bag and I bought my first PC. WordFast was the result of a dare, if you like. Someone told me he was writing macros using Microsoft's macro language, VBA. This guy said that a translation memory engine could never be written using MS Word's macro language because it was just too slow. I just said, 'no' and began to write a sort of rough sketch of a translation engine using the resources of Microsoft Word. That was back in August 1999 and, after a week of writing routines, I actually found that it worked. Of course, VBA is a slow language in itself if you compare it to C or Pascal or Delphi, but speed has nothing to do with the language of the platform itself, it has to do with the way that you write your software. I discovered ways of executing a database search extremely fast. By September 1999 I had a translation memory engine that ran as fast as any other but only used Microsoft Word as a

platform. Typically, a translation memory engine has to scan or search a huge database that can be hundreds or thousands of megabytes, looking for an exact match, which is no big deal, but also looking for a fuzzy match. You have a rough idea of what you are looking for but the software has to pinpoint the location of something that approaches your search. That is extremely difficult. I actually wrote the necessary algorithm that can perform this very fast just using Microsoft Word's macro language. That was a breakthrough. I could use the first sketch for version 1 of WordFast. The format of the translation memory is an open format. You could take a WordFast translation memory and open it with Excel, Word or Access, etc. It's not a proprietary format. That applies to the glossaries as well as the translation memory. It's all pure text."

Can this be true? Just how open is this open system? Can anyone take the lid off and customize its behavior? "No, it's not an open source project," says Champollion.

The format of the translation memory is an open format. You could take a WordFast translation memory and open it with Excel, Word or Access, etc.

"The source is mine but I've made what I call entry points, which means that any programmer could ask WordFast to segment a document, to find a match in a translation memory. In fact, all the basic tasks of the translation memory engine can be accessed externally. Besides the manual, there is what I call the White Book, which is more technical and shows programmers how they can write applications to enable them to use the WordFast translation memory engine."

And what about that inevitable question, 'Isn't this yet another TRADOS clone?' "All products with a segment-based translation memory will inevitably look alike," replies Champollion. "I use colors to represent match types. I use green for exact matches, there is no copyright on green. A green light means, 'It's OK, move on.' Yellow or orange means 'caution'. I use a sort of grayish background to represent a no-match. I don't think this breaks any rules. Many other products use colors the same way. My shortcuts are not identical. There is some overlap with TRADOS but

Photo: Seiji Champollion

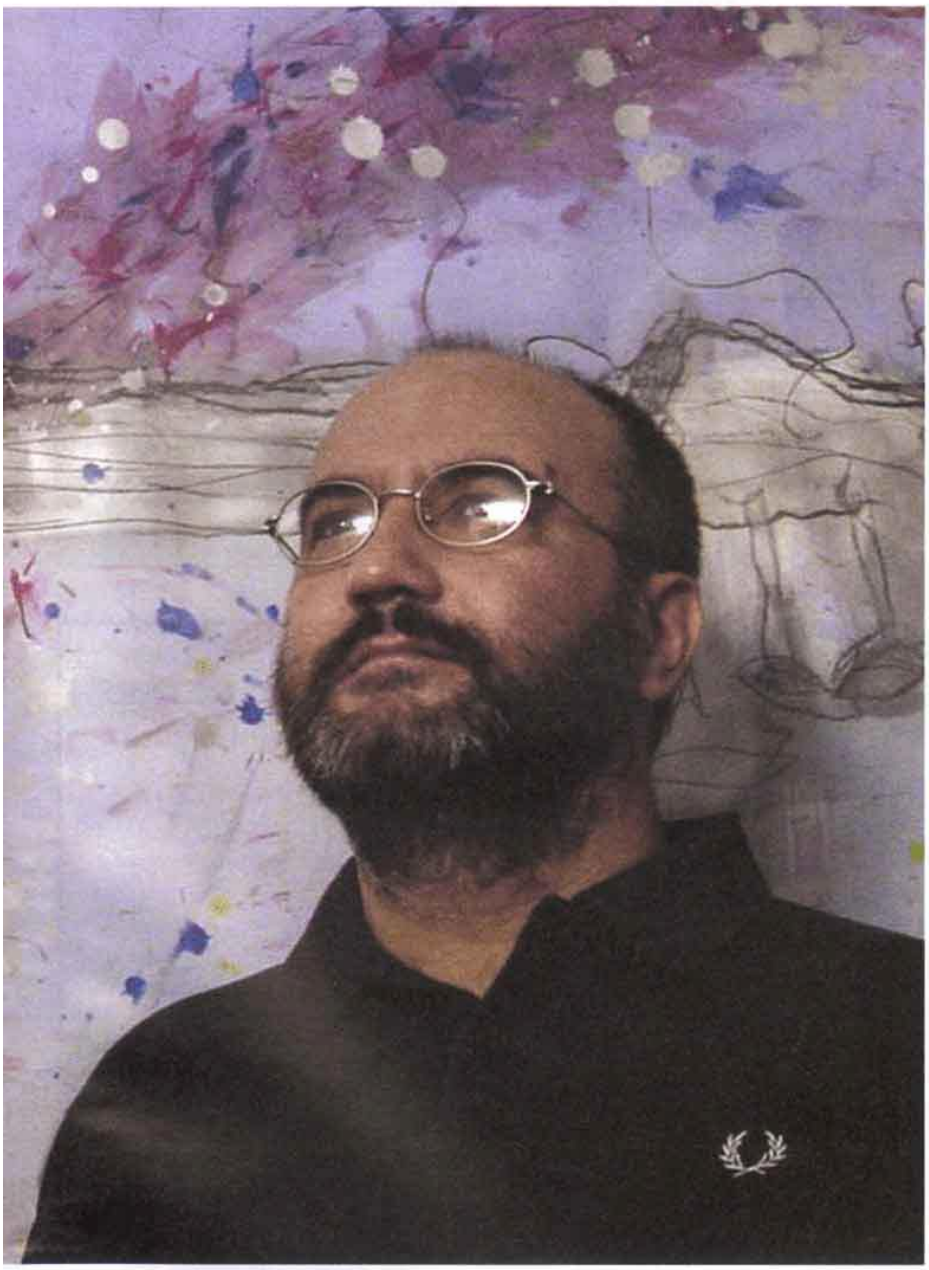
others are different. In any case, there are quite a few features in WordFast that TRADOS doesn't have, so you couldn't really call it a clone. I built on a concept that existed long before TRADOS came along. Look at word processors. They all look and feel the same. WordFast also has a totally different terminology approach. In addition to its built-in glossaries, WordFast can be interfaced with any third party dictionary. Translation memories can also be shared through a network. There is no complicated setup to do this. You open the same translation memory and that's it. From that point, WordFast knows that someone else is using the same translation memory. It's just like sharing an Access file. Up to 20 people can share a translation memory over a network and a translator's input is automatically shared by all the others on the same LAN. Web-based sharing will be done in the next version. The other concern that has been addressed is interchange with other translation memory software. WordFast can read IBM TranslationManager .exp files. It can read TMX files, which means it can read practically any memory from any tool. It also reads TRADOS native .tmw files under certain conditions, and text exports. The translation memory of WordFast is so simple that you can open it with Excel. You could paste segments into that database and resave it as a text file and it works. It is a very transparent and open format. In the latest release, the translation memory can be maintained in Unicode format or simple text format. WordFast will automatically detect a Unicode translation memory."

If you want to enter the entire US Government terminology list, if you want to have one gigabyte of glossary on line, WordFast will open that in no time and quickly find an entry.

So far, so good. What are these other features? "We offer a suite of tools, called PlusTools," says Champollion. "Initially, PlusTools was mainly geared for doing search and replace among vast numbers of files. You know, if you have a project with 200 files and the customer comes and says, 'This has to be replaced by that'. You have to open up every file and do it by hand. So this can batch-process up to 1000 files in one go, with all the refinements that Word offers. There is a tagging utility that allows you to tag and then un-tag HTML, SGML and XML files. This prepares files for translation and then reconstructs the files after translation. There is a conversion utility. You can convert files from one Word format to another in batch mode. You could also set up a conversion table to convert, for example, non-Unicode Russian text to Russian Unicode text. There is also an alignment utility to align existing file pairs to create translation memory. This will produce a WordFast translation memory but WordFast can export this to TMX format. So, you could use this tool to align files and export the output to any translation memory software. There is also a glossary that has just been ported to Unicode, containing CJK characters, that complements the existing WordFast glossaries. The WordFast glossary format is tab-delimited pure text, whether it's Unicode or not. If you have an Excel file with a glossary set up in columns and you save that as a text file, you have a WordFast glossary file. You could also have it in an MS Word table and save it, which would produce the same sort of file. At the moment, the established fields are Creation Date, Creator Identity, then you can

Photo: Courtesy B. Vacon

add fields like Client, Subject and so on. You can then merge glossaries or extract sub-glossaries from the main one. The technology that I developed for the translation memory has been reused in the glossary. There is no limitation of file size. If you want to enter the entire US Government terminology list, if you want to have one gigabyte of glossary on line,



WordFast will open that in no time and quickly find an entry. That sounds amazing but that's the way it is. There is also a built-in terminology compliance feature. You enter the customer's terminology in WordFast using wildcards so the system recognizes all forms of the same term. As the translator moves from one segment to the next, WordFast will scan the source segment and, if it finds a customer's term in the source text, it will expect to find the preferred target term in the target segment. If that is not the case, the translator is warned that there is a discrepancy. The translator then has the opportunity to change the target term, or not, and then move on. This process can also be done at a later stage in batch mode and a report is produced. This approach not only applies to terminology but also typographical rules and non-translatable elements, for example, numeric parameters and tags, if it is a tagged file."

So, we have an apparently fully-operational translation memory system, but why is it *free*? "The main reason is that Logos noticed WordFast, and expressed interest in it. They offered not only to share their own technological developments in the same area, but to sponsor WordFast so that it is made available for free to the translation community. Although a later version could possibly go commercial, the basic version will most likely remain free. When I first made it available, there were about 20 units per week being downloaded. Now it's about 40 per day and we have well over 1000 users, so Logos was right in deciding to keep WordFast free. The user base is also contributing to technical support. We have a very active email-based user group constantly exchanging information and supplying feedback. A nobler, less commercial motive was to make the technology freely available to those parts of the world where the cost of currently available software makes adoption impossible. They have the need but not the necessary financial resources. Of course, it would also benefit the universities training future translators."

When I first made it available, there were about 20 units per week being downloaded. Now it's about 40 per day and we have well over 1000 users.

So, does it work? Only you can say. Fortunately, it won't cost you anything to find out. Healthy cynicism is always wise when investing in new tools. However, I would have thought that people should be encouraged let their imaginations run wild and be adventurous. That's how innovation comes about and, hopefully, stagnation is eroded. Maybe, just maybe, Yves Champollion and people like him are showing the way. Still, not everybody has an ancestor that cracked the Rosetta Stone code. The good news is that the Logos group has been impressed enough to underwrite the future of the software and their developer, Bruno Vaccari, is collaborating with Champollion on the next version of WordFast.

WordFast can be downloaded from http://champollion.net.