

## Mind the gap

by Juliet Macan

### Abstract

People movers, escalators and high-speed trains are the technology used today to automate mass people transit. For these forms of transfer there are strict safety regulations and legal aspects which make it imperative to deal with the problems that can arise at the interchange or "crossover" points between the departure and arrival platforms and the moving section in between. They carry warning signs to avert those in transit of possible dangers. I think that there are some analogies we can trace with the transfer of words between various software platforms involved in word processing, content management, web sites and translation tools. Maybe here there should also be warnings!

The enormous increase in volumes of text to be transferred between different languages and ever shorter deadlines require a new approach on the part of those planning and writing the original texts, as well as those involved in its transfer to other languages, not to mention those who receive and use such texts. Greater cooperation between the various "stations" could help to refine the technologies adopted for the benefit of all those involved.

Translation itself is a complex task, just like walking or climbing stairs, but the introduction of automation for the creation of text, its processing and translation adds new variables which must be considered to avoid cancelling the expected benefits. The increasing number of destinations/languages involved, the sophisticated vehicles/software and chronic shortage of time and resources makes careful planning essential. Indeed, a crucial part is the planning and building of the original departure platform, taking into account the route/routes to be travelled. This should include awareness of problems linked to the type of vehicle - compatibility with TM tools - and mapping - style-sheets "converted" into other "locales", etc. This knowledge can be further refined by careful networking with experienced LSP's who can provide precious indications for a smoother

transit, offering advice on the basic materials - terminology - as well as the "local currencies" - language specific problems.

Finally, on the arrival platform, the accuracy of the transfer system should be vetted by the reviser/local check who should provide feedback on the efficiency of the transfer and point out any technical problems that need sorting before the content is made available to the final end-user. Currently, there are often serious misalignments between the various stages which can only be eliminated by closer cooperation. It is important to adopt a blue-print right from the start. Furthermore, this blue-print could also help the software companies to provide better tools all along the line and thus facilitate a smooth and accurate transfer.

### **Introduction**

Today I would like to talk about some of the gaps encountered with the transfer of words between various languages and also between different software platforms involved in word processing, content management, web sites and translation tools. The enormous increase in volumes of language content to be transferred between different countries and ever shorter deadlines requires a new approach on the part of those planning and writing the original content, as well as those involved in its transfer. Furthermore the technology available to facilitate this process must be carefully tuned to ensure maximum efficiency.

### **Destination**

It is essential to determine the precise purpose and destination of the content before departure. This may appear obvious but unfortunately this is not always the case and can result in disappointing results. When ordering a translation the customer does not always remember that it is not sufficient to indicate a language without specifying the destination country. It is a bit like going to the United States with your cell phone, without previously checking whether it is tri-band. This gap is frequently underestimated. But if we take Europe for example, and the German language, there are some quite substantial differences between the German used in Germany, Austria and Switzerland.

Now let's examine some different "destinations".

Individuals wishing to communicate in another language or travel abroad for pleasure, study, medical or other reasons may require visas, letters, certificates, etc. requiring translation in another language. These documents can be highly personalized and therefore not suitable for automatic processing. Thus another "gap" between one location and another. A bit like electrical plugs and sockets.

If we move to the business world, where it is possible to identify perhaps five main types of content: legal/financial, sales/marketing, software, installation and user instructions, product support, these gaps take on another dimension. Here again it is essential to define all the destinations and prepare a schedule based on consultation with the vectors. If well planned, most of this content can benefit from automation to reduce the time and cost of the transfer and most important - improve efficiency. But despite some of the promises made by some companies selling Translation processing tools, the efficiency of automation depends heavily on the preliminary authoring stages, preparation of terminology and reference material, the "transferability" of the originals and the universal applicability of any style sheets selected to bridge or smooth the gaps.

### **Words are slippery**

Since we are dealing with languages, I think everyone here will agree that the passport is **terminology**. I think we need an enormous billboard here, not just a warning sign!

A company should establish the basic terminology it uses, especially with reference to its products, ideally with the relative definitions, so that everyone in the company uses the same vocabulary. This makes communication easier and eliminates misunderstandings. It also greatly facilitates an accurate transfer to other languages. The larger the company the more important this becomes since it is unlikely that the legal department will have much contact with the installation technicians or marketing with product support, let alone the offices in Nebraska with those in Shanghai, but they all generate content which requires transfer to

multiple destinations. Any ambiguities in the original content will be magnified and multiplied in the target languages.

### **Vehicle**

Then, before the content is embarked, it is necessary to check not just the destination but also its compatibility with travel. Whether it a MS Word or PowerPoint document, a Framemaker or InDesign file, or in HTML, XML or exported from a CM system? Here there can be a series of possible hitches. First of all is the vehicle appropriate for a smooth journey. What you see is not always what you get! A case in point is a PDF - most of us are well aware of the problems with processing elaborate PDF files when the original is not available.

And then the target language(s). Has space been allowed for the expansion of "bulkier" languages? Is the vehicle Unicode compliant? Is the original a multilingual document, with the different languages in separate layers, paragraphs, columns or generally mixed together?

Another gap which should not be underestimated is whether the vector/translator(s) has the "vehicle/version" or "driving experience" required to deal with the complete transfer. The vector may possess MS Word, but is it the right version? For many reasons it is preferable for a translator not to have to deal with DTP programs. Another justification for the increasing recourse to conversion with TeX tools, outside MS Word, to overcome these difficulties as well as reduce time and cost. This means however that the author should be aware that the content will be exported/imported from the original format for processing and therefore that smooth "text-flow" is of utmost importance. Any phrases or sentences split by hard returns, for layout purposes, will be difficult to handle and impossible to automate. The same applies to text in graphic form, revisions and comments which may stop the transfer "dead in its tracks". The author should also be aware of the importance of consistency and again the use of standard terminology, since the translator will be mother-tongue of the target language; if a technical translator encounters the terms "Warning sign" and "Safety sign" in the same document, it is legitimate to wonder whether they refer to different types of signs!

## **Vectors**

Then we have the vectors. There are the "back-packers" or do-it-yourself vectors aided by web services and machine translation where the onus is on the user to insert unambiguous text and verify the correct transposition. Still a road scattered with potholes, but useful for gisting. Generally this type of vector is adopted to transfer content that would not otherwise leave the ground. But here again, an intelligent investment in terminology can provide interesting results.

On the other hand, an in-house translation department can have a privileged position as there is generally a more direct link with the author of content, and also with local branches of the company or agents who can be consulted. This department is normally responsible for the preparation and maintenance of company terminology and therefore probably the most efficient "conveyor belt", but not all companies can afford or have enough work for an in-house department. Unfortunately, as Renato Beninato, from Common Sense Advisory, once pointed out, in the budget of a large international company, the sum spent on translation often comes well below that spent on lavatory paper and thus the "visibility" and importance attributed to it by the all-powerful Purchase Department. It is a question of economics, or perhaps those providing the technology or the "transfer" services are using the wrong arguments! Another kind of communication gap.

Then we have the single vectors/freelance translator, frequently called in to deal with the "overflow" but without all the advantages of in-house staff. Thus a few gaps may appear in the conveyor belt since even an experienced freelancer, working for a variety of customers on different content, often with tight deadlines and little specific training or feedback for each project, will have to deal single-handed with any technical and linguistic problems that arise. The former can unfortunately interfere with the attention dedicated to the latter.

A more solid alternative can be offered by Language Service Providers, both monolingual and multilingual vendors. These companies, like shipping agents, receive the content from their customers together with the relative instructions. They analyse the content to identify any possible problems,

before processing. This is perhaps the most crucial part of the "journey". It is here that any gaps need to be identified and dealt with, to avoid hold-ups during the transfer stage or on "disembarking" at the final destination.

### **Technology**

The results obtained with TeN tools can be astounding, but they depend heavily on the quality of the source materials. Unfortunately, although these bring some problems to the surface - such as hard returns in the middle of a sentence - these are generally not detected automatically in the preparation stage. I was very pleased to note last year that Trados 7 was able to flag an MS Word document containing revisions, although more recently I heard that a translator actually complained about this. There again I do admit that I find Trados 8's sensitivity to irregularities in Adobe FrameMaker files somewhat frustrating, but it comes up trumps with structured files. The main problem is that the smooth operation of translation tools is based on the presumption that the original content is correctly prepared. A bit like an airline presuming that all the adult passengers are the same size and shape, and understand the language spoken by the steward.

To my knowledge there are only two QA tools currently available that can check monolingual source material: Acrocheck (by [www.acrolinx.com](http://www.acrolinx.com)), an impressive but expensive corporate tool (for German, English, French and Spanish) and the Formatcheckers (by [www.star.com](http://www.star.com)) for MS Word and Framemaker. All the other QA tools check the translated document - against the source, but this makes little sense if the latter is imperfect. So, technology can be a mixed blessing. It can also be a nightmare! Despite the increasing use of CM systems, with the precise purpose of standardising content, there are still many gaps to be filled. Moreover, the use of authoring tools is still far from widespread. If these technological solutions could be combined with **clearly defined terminology** we could smooth out many of the gaps to everyone's advantage.

## **Other gaps**

It is also important to remember that there are other headaches such as character encoding and font mapping that need to be dealt with and require expert hands to identify the various gaps before departure and devise adequate solutions to ensure a smooth journey. Once the gaps have been plugged the travel package is assembled and dispatched for translation. This should include specific instructions concerning the purpose, reference material, terminology. During transfer, further support should also be available to answer queries, provide feedback and draft PDFs for revision, run QAs and validation tests.

On arrival, the content can still encounter obstacles during export, testing and reorganization in the destination format. These gaps are not just linked to the format, but can be complicated by the "local currency" or character encoding. A good example was provided by Maltese in Microsoft programs and in Adobe Frame Maker in 2004 when this language became one of the EU languages. And obviously it is also possible to trip up on "customs and traditions" or locale if the destination country is not clearly indicated from the beginning. British and American English provide many examples.

A couple of years ago I mentioned the unfortunate results of a Trados marketing message to some big international translation customers, i.e. the suggestion that 100% matches merely recycled translation with the inference that they should therefore be free. This created great and false expectations, and did not help to create a cooperative relationship with the vectors. Other attempts to short-circuit the process have been made by customer who have purchased all-in-one tools like Across and naturally expect to just enter the content in one-end and wait for the translation to come out at the other. At the end of the day, although I have to admit that I am a technology geek, I find that the automatic, web-based, project management systems available today remind me a bit of trying to run a Ferrari along the Italian roads from Rome to Reggio Calabria and expecting the same results obtained as on the race track at Maranello with all the mechanics-in-waiting.

So in my opinion, the critical points for a smooth and easy transfer are:

- planning and preparation;
- assessment of travel-worthiness and the condition of the vehicle:  
compatibility of source content and format with language processing;
- portfolio of experienced and efficient vectors;
- technology but most of all terminology management

But you won't get very far without a passport.