Computer-aided technical translation workflows - man-machine in the construction and transfer of corporate knowledge

Elia Yuste (Zürich)

Abstract

The translation of technical documents is no longer a process to be tackled once the source text has been produced elsewhere in the corporation or institution. Instead, translation and intercultural issues are taken into account prior to the production of the technical source-text which, most frequently, gains a dynamic feature. From this standpoint, one might think that the translator loses the leading role against the technical writer, who would become the key intercultural communicator for the global marketplace. However, as it has happened to other stakeholders in the multilingual documentation chain, the translator is no longer empty-handed and is gradually developing some technologically-driven competencies that allow them to expand their battle area and contribute to knowledge transfer in our society. This paper provides an overview of the latest workflow patterns leading to multilingual technical documentation production and management, their ideal circumstances and problems, as well as deriving consequences for higher education degree programmes in translation and technical communication.

1   Introduction

Large and middle-sized corporations that have decided to go multilingual in the last few years have shifted or are in the process of shifting to a major operational development as far as the production and management of their multilingual documentation is concerned.   This consists of pursuing the goal of multilingualism at an early stage within the production chain, therefore acknowledging the cyclic and multi-componential nature of such production. Until just about a decade ago it was standard practice to write the corporation's monolingual technical documents in an isolated technical writing department (perhaps by a group of experts with no particular language flair) and then have the resulting texts passed on to a language support team, in-house or elsewhere, who would traditionally be conceived as the solely responsible for the corporation's multilingualism effort and success. More recently, however, authors such as Hartley/Paris (1997) point to the fact that the monolingual technical documentation writing, followed by its translation into various languages, is being replaced by new documentation production workflows, whereby a technical writing team is in charge of delivering

technical documentation in several languages. If we do not explore that idea further, both technical translators and translation training institutions could begin to worry. They might be inclined to think that translators are being somewhat excluded, no matter how well equipped and specialised in a particular technical domain they might be.

It thus seems necessary to examine these innovative multilingual technical documentation production workflows and to find out what steps and agents are involved therein, and what techniques and abilities are to be owned by technical translators in order to not only remain but also shine in the multilingual communication game.

2 Procedures, strategies and trends

2.1 Pre-editing is as important as post-editing

When a company or institution handle technical documentation without having taken into consideration translation, intercultural and technological issues at the right time, that is, at an early stage and on an on-going basis, they may face serious difficulties that also come at a cost: ambiguities and inconsistencies, deficient translation memories, or expensive and time-consuming post-editing. Even if some of these may still be found, more often than it would be desired, thus leading to potentially damaging effects for the institutional or corporate image, interestingly enough, there is a growing tendency to reduce the post-editing phase as much as possible and to optimise the multilingual documentation production chain from its starting point, that is, a pre-editing, preliminary or preparatory stage.

As Göpferich (2000: 2) puts it, this can be achieved if the so-called pre-editing is based upon two essential elements:

- "die Zusammenstellung neuer Texte aus bereits existierenden (und übersetzten) Textbausteinen [und]
- den Einsatz kontrollierter Sprachen"

In other words, pre-editing would be based on the building up of new texts from textual strings of existing versions (i.e. a sort of 'writing memories'), both monolingual and translated, together with the implementation of controlled languages.

Both procedures are interesting for two main reasons: on the one hand, language resources are reused to spare time and effort (an idea that has already consolidated through the widespread use of commercial translation memory systems) and now already from the writing stage (Allen 2000); on the other hand, the usage of a controlled language is proposed. Controlled languages are gradually becoming more popular as they have proved to be useful in combination with machine translation (MT) (Allen 1999) and for enhancing other computational linguistics applications.

On the whole we have been experiencing a turn to implementing pre-editing measures in the dynamic multilingual text production cycle. In the next subsection, we present a currently popular framework where this is precisely the case.

2.2 The GILT framework

The increasingly wide adoption of the GILT (an acronym that stands for Globalisation, Internationalisation, Localisation and Translation) framework explains why the structured development of multilingual contents is to be sought from the beginning. But let us explore why and how this can be achieved.

There are plenty of references (both academic and from industry specialised associations, such as LISA) and web sites aimed at describing in detail this framework, its philosophy, components, ins and outs, originally from an economic standpoint, then significantly diversifying itself. Even if we are primarily interested in the communicative and intercultural aspects of GILT, it goes without saying that globalisation mechanisms have brought about an unprecedented consideration of multilingual matters.

Many multinational companies willing to have a truly global presence have become aware that imposing the dominant English language does not generate enough profit. As a result global players are keen to take into account important issues of internationalisation and localisation, such as "Will the graphics and colours in our website suit our users (and potential buyers) in Brazil, Korea or Greece? How do we go around to make our product look as if it was actually originated at X place so that the users of X's language and culture like it?"

A corporation's globalisation strategy entails the need to develop standard procedures that allow for the creation, retrieval and exchange of corporate content units among several interconnected applications. Internationalisation measures come then into place. What at the beginning may seem to be too complex enables a better content scaffolding and classification. The corporate knowledge architecture is then further enriched by the localisation stage, aimed at pursuing software interoperability and an affinity of digital and network corporate content as well as documentation for each *locale*. Since these highly technological processes have been (and are being) devoted much attention, some question the role of the traditional translator and even if there is still room for human translation in the GILT approach.

Our standpoint is far from pessimistic. As a matter of fact, we have waited for over a decade to corroborate what we already knew: a culturally informed translation competence is indispensable to transform any content into a fully working multilingual entity; adjusting it to each final user to ensure that the corporate knowledge is properly transferred. Nonetheless it will be important to examine how that competence has evolved and keeps evolving.

2.3 Evolution of the culturally informed translation competence

In this subsection we will look at how this competence is evolving, also taking into account what has been exposed in section 2.1. Let us imagine the following scenario: a source text is written in a *controlled* manner to be processed by a machine translation system, which then produces a multilingual output in several languages. It will be the system which will somehow have a translation "competence", yet the translator will not be out of the game.

In many other areas of our daily life we have already got used to the fact that technological advances save us a great deal of time and avoid task duplication. In other words, the man-machine relation is based upon complementary rather than opposite terms. By the same token the competence of skilled human translators is and will remain necessary - what may have evolved are the action lines or work patterns within more dynamic workflows and cyclic processes, such as that of the multilingual documentation production. In this setting the human translator's competence evolves as to feature in a pool of technologically flavoured linguistic tasks. Let us just name a few:

- translation of some text types, such as marketing or publicity texts, unsuitable for machine translation;

- preparation and follow-up of CAT (computer-aided translation) projects (including terminography work, memories management and updating, etc.);

- CAT tools' maintenance and enhancement (including not only well-known commercial translation memory systems but also MT applications; the language professional will be in charge of "feeding" these systems to guarantee the best possible results);

- design and improvement of controlled languages and their corresponding writing-aid tools (i.e. conformance checkers);

- sophisticated localisation projects (including careful assessment of cultural issues).

Moreover, the attempts to incorporate specialised human knowledge in corporate knowledge bases (as described by Hartley/Paris 1997: 113) in order to automatically generate multilingual text make even clearer that human linguists are the ones designing the rules that the machine subsequently applies. It is then evident that man-machine competencies are expected to co-function rather than to replace one another.

3   Ideal processes … without losing one's foothold

3.1   "It is not always easy, not everybody knows..."

Interesting though the automatic generation of multilingual text can be, the commonest processes of multilingual documentation production are not yet so *advanced*. This kind of move is often perceived as one of various process optimisation options, with a potential that is not yet crystal clear. However, what does already constitute an ideal process (and so it is being taken for) is the one that we describe next, mainly characterised by being a non-stop cycle with several components and agents. This ideal process could be summarised as follows:

a) First, a controlled language is put in place to help create the corporate documentation. It relies upon a consistent vocabulary range (e.g. a corporation-wide agreed repository of terms) and a careful syntactic rule set, with the corresponding conformance checkers. There could also be a writing memory tool.

Elia Yuste: Computer-aided technical translation workflows -
man-machine in the construction and transfer of corporate knowledge

71

b) These texts would then be imported to a translation memory system containing stored translated segments to be retrieved as appropriate. Those segments for which identical o very similar matches have been found will also be entered in the memory.

c) The segments without a match are then processed by a machine translation system under constant supervision of a human linguist, who also takes care of the system's updating and output post-editing, if necessary.

There would also be room for plenty of other important activities relating to *quality control*, for example: the translation itself or the process components/resources should be validated by experts (e.g. validation of domain terms in a term-base compilation); the project course and results is to be evaluated not only by the project manager but by other agents inside (e.g. a department manager) or outside (e.g. the client or the project commissioner) the corporation, etc. On the basis of the results obtained some stages or steps in the process would have to be modified or improved, no matter how early in the process that could be. This is why we wish to underline - once more - the cyclic nature of a multilingual documentation production workflow nowadays.

One cannot deny the importance of initial stages of such workflow. And so it is useful to stress that there are a number of preparatory tasks where the translator or corporate linguist can already intervene. To name but a few:

- Term extraction - this can be one word or a multi-word unit (for which special software applications and different types of corpora can be used);

- Definition of terminological work guidelines;

- Monolingual and multilingual terminography (usually computer-aided);

- Definition of templates, format directions and standards to manage, retrieve and leverage corporate content;

- Creation of style guidelines and controlled language rules to facilitate and systematise the writing of corporate documentation at monolingual level (e.g. elimination of ambiguous wording, consistent use of corporate language keywords, etc.); standardisation of phrases, slogans, acronyms, definitions, etc. (of particular importance in very technical texts);

- Dynamic and detailed management of reference resources for process components, such as terminology extractors, translation memories and machine translation systems.

3.2 Challenges for the translator and the translator trainer, today and tomorrow

As we have just observed, the tasks and functions of both translator and technical writer get assimilated, are shared and reinforced with one another. The reason behind this is that they have a common objective, namely, to transfer corporate knowledge to the rest of the world. The fact that translators and technical writers work in a convergent and technologically driven fashion represents a major challenge for universities and specialised training institutions, who are expected to update their degrees and specialisation courses on a regular basis. Far from the

not so recent past where translation and technical writing belonged to two worlds apart, we propose a university degree or postgraduate course syllabus that, under the title "intercultural technical communication", reflects and adapts to the multilingual and intercultural needs of a global market and the knowledge society.

For this purpose the following areas would have to be taken into consideration:

- Foundations of technical communication (textual production in students' mother tongue, document structure standardisation - concept, standards, procedures, usage of tools and language resources, etc.);

- Cultural-linguistic flair in a foreign language - L2 (with option to a second foreign one, L3) and translation competence;

- Field or domain specialisation (including also the mastering of computational linguistics techniques to be inserted naturally in the training assignments).

Our future *intercultural technical communicators* would study the first three academic years at their home university. This would then be followed by a fourth year at a foreign university department where their L2 is spoken. This stay abroad would include an extra work placement at a company or institution, about which a final report would be submitted. After that, they could study, at their home university or abroad, a postgraduate degree aimed at specialising in one or more of these subjects: computational linguistics, sociolinguistics and intercultural studies, multimedia programming, a third foreign language, etc.

Little or no benefit will be derived from the efforts to set up these innovative, interdisciplinary degree programmes unless new forms of collaborative teaching and learning (Kiraly 2000) are implemented. Some authors argue that assignments should be project like, simulating and getting nearer to the real world, solving problems (Palomares 2003) and observing current production and management processes of multilingual communication (Yuste 2002b), without losing track of technological advancements, as study objective (Montalt 2003: 326) and a working tool (Tebé 2003, among others). Let us stress that the main goal here is to learn how to construct knowledge together in order to be able to transfer it later when the communication act or form takes place.

4   Conclusions

The human translator, as a corporate knowledge player, is and will remain essential for the foreseeable future. The new technologies surrounding them make their work more consistent and reliable, supporting instead of excluding them, as M. Kay predicted in 1980. Furthermore, the translator is to be broadly conceived as an active corporate linguist that is capable of improving and interacting with several applications and tools. Through the creation, maintenance and leverage of corporate language resources they will also develop and enrich their competence on an ongoing basis. In practical terms, they will become indispensable agents of the knowledge society. This must be noted not only by future technical translators but by their trainers. To provide orientation in that respect, it has been argued that the new technologies are to be both a learning/teaching support and a subject of study. This also requires new

Elia Yuste: Computer-aided technical translation workflows -
man-machine in the construction and transfer of corporate knowledge

73

learning/teaching methods that echo the reality and needs of the global communication market.

References

Alcina Caudet, Amparo/Gamero Pérez, Silvia (eds.) (2002): *La traducción científico-técnica y la terminología en la sociedad de la información*. Castelló de la Plana, Spain. (= *Estudis sobre la traducció* 10).

Allen, Jeffrey (1999): "Implementing Controlled Language and Machine Translation in the automotive industry". Invited Keynote at the *Multilingual Documentation for the Automotive Industry TOPTEC Symposium*, sponsored by the *Society of Automotive Engineers (SAE)*, Localisation Industry Standards Association (LISA) and ALPNET. 21st -22nd October 1999. Amsterdam.

Allen, Jeffrey (2000): "Adapting the concept of 'Translation Memory' to 'Authoring Memory' for a Controlled Language writing environment". *Translating and the computer 21: Proceedings of the twenty-first International Conference on 'Translating and the Computer', 10-11 November 1999, London* (unnumbered).

Colominas, Carme (2003): "Los corpora como herramientas de traducción". In: García Peinado, M. Á./Ortega Arjonilla, E. (eds.): *Panorama actual de la investigación en traducción e interpretación*. Vol. 2. Section III - Localización, traducción asistida y traducción automática (aplicaciones didácticas y profesionales). Granada, Spain: 363-372.

Göpferich, Susanne (2000): "Der Technische Redakteur als Global Player". *Dokumentation erstellen - übersetzen - managen Online* 11.5.
http://www.doku.net/artikel/dertechnis.htm [last consulted in January 2005]

Hartley, Anthony F./Paris, Cecile L. (1997): "Multilingual Document Production: From Support for Translating to Support for Authoring". *Machine Translation* 12, 1-2: 109-129.

ISG (Intercultural Studies Group) (2003): *Localization and Translator Training. An Online Conference*.
http://isg.urv.es/seminars/2003_localization_online/ [last consulted in February 2004]

Kay, Martin (1980): "The Proper Place of Men and Machines in Language Translation". Republished again in 1997 by the original author at *Machine Translation* 12: 3-23 and in 2003 by Nirenburg, S./Somers, H. /Wilcks Y. (eds.): *Readings in Machine Translation*. Cambridge, Massachusetts/London: 233-238.

Kiraly, Donald C. (2000): *A Social Constructivist Approach to Translator Education: Empowerment from Theory to Practice*. Manchester.

Mata Pastor, Manuel (2003): "Algunas consideraciones básicas sobre la investigación en el ámbito de la tradumática". In: García Peinado, M. Á./Ortega Arjonilla, E. (eds.): *Panorama actual de la investigación en traducción e interpretación*. Vol. 2. Section III - Localización, traducción asistida y traducción automática (aplicaciones didácticas y profesionales). Granada, Spain: 289-312.

Mayer, Felix/Schmitz, Klaus-Dirk/Zeumer, Jutta (eds.) (2004): *Terminologie und Wissensmanagement. Akten des Symposions "Deutscher Terminologie Tag e.V."*. *Cologne. 24th-27th March 2004*. Köln.

Montalt i Resurrecció, Vinzent (2003): "La traducción de géneros electrónicos: el caso de la localización". In: García Peinado, M. Á./Ortega Arjonilla, E. (eds.): *Panorama actual de la investigación en traducción e interpretación*. Vol. 2. Section III - Localización, traducción asistida y traducción automática (aplicaciones didácticas y profesionales). Granada, Spain: 313-328.

Multilingual Computing and Technology (2003): *Writing for Translation - The Guide*. Supplement #59: October/November 2003.
http://www.multilingual.com/FMPro?-DB=vendors&-lay=CGI&-format=gettingStarted/welcomeGettingStarted.htm&-find [last consulted in January 2005].

Nyberg, Eric/Mitamura, Teruko/Huijsen, Willem-Olaf (2003): "Controlled language for authoring and translation". Somers, H. (ed.): *Computers and Translation - A Translator's Guide*. Amsterdam/Philadelphia: 245-281. (= *Benjamins Translation Library* 35).

O'Hagan, Minako/Ashworth, David (2002): *Translation-Mediated Communication in a Digital World - Facing the Challenges of Globalization and Localization*. Clevedon, England. (= *Topics in Translation* 23).

Palomares Perraut, Rocío (2003): "Premisas cognitivas del proceso documental del traductor". In: García Peinado, M. Á./Ortega Arjonilla, E. (eds.): *Panorama actual de la investigación en traducción e interpretación*. Vol. 2. Section III - Localización, traducción asistida y traducción automática (aplicaciones didácticas y profesionales). Granada, Spain: 399-411.

Sprung, Robert C. (ed.) (2000): *Translating into Success. Cutting-edge strategies for going multilingual in a global age*. Amsterdam/Philadelphia.

Tebé, Carles (2003): "Las herramientas informáticas en la didáctica de la traducción especializada". In: García Peinado, M. Á./Ortega Arjonilla, E. (eds.): *Panorama actual de la investigación en traducción e interpretación*. Vol. 2. Section III - Localización, traducción asistida y traducción automática (aplicaciones didácticas y profesionales). Granada, Spain: 329-342.

Wright, Sue Ellen: "The InfoCycle: Information Feedback for Interactive Multilingual Document Production".
http://appling.kent.edu/ResourcePages/LTStandards/Chart/infocycle.html [last consulted in January 2005].

Yuste, Elia (2001): "Making MT commonplace in translation training curricula: too many misconceptions, so much potential!". In: Maegaard, B. (ed.): *MT Summit VIII*. CD-ROM *Proceedings of the 8th International Machine Translation Summit - 'Machine Translation in the Information Age.' Santiago de Compostela (Spain), 18th-22nd September 2001*. Geneva: 45-49.
http://www.dlsi.ua.es/tmt/proceedings.html [last consulted in January 2005].

Elia Yuste: Computer-aided technical translation workflows -
man-machine in the construction and transfer of corporate knowledge

75

Yuste, Elia (2002a): "Language Resources and the Language Professional". *Proceedings of the First International Workshop in Language Resources for Translation Work and Research. LREC 2002, 28th May 2002. Las Palmas de Gran Canaria (Spain).*
http://www.ifi.unizh.ch/cl/yuste/postworkshop/repository/eyuste.pdf [last consulted in January 2005].

Yuste, Elia (2002b): "MT and the Swiss language service providers: an analysis and training perspective". *Proceedings of the II International Workshop on Teaching Machine Translation (TMT).* UMIST, Manchester. 14th-15th November 2002. Organised by *EAMT* and the *Natural Language Translation Specialist Group* of the *British Computing Society (BCS)*: 23-32.
http://www.co.umist.ac.uk/~harold/teachingMT/ [last consulted in January 2005].

Yuste, Elia/Cerrella Bauer, Silvia (2004a) [In print]: "Circumstances, challenges and consequences of implementing a quality-geared and technology-aided process of translating: a case study". 90-minute workshop at the *IV Jornadas sobre la Formación y Profesión del Traductor e Intérprete.* UEM, Madrid, 25th-27th February 2004. CD-ROM Proceedings due to appear in spring 2005.

Yuste, Elia/Cerrella Bauer, Silvia (2004b): "Implementing a quality-geared and technology-aided process of translating: a case study". *Bulletin of the DÜV* (Agentur der Dolmetscher- und Übersetzervereinigung), May 2004. Switzerland: 12-18.

Yuste, Elia/Cerrella Bauer, Silvia (2004c): "Circumstances, challenges and consequences of a quality-geared and technology-aided process of translating: a case study". *Hieronymus*, the professional quarterly journal of the ASTTI (Association suisse des traducteurs, terminologues et interprètes).