TRANSLATION TECHNOLOGIES: SCOPE, TOOLS AND RESOURCES

Amparo Alcina

Universitat Jaume I

Amparo Alcina

Departamento de Traducción y Comunicación

Facultad de Ciencias Humanas y Sociales

Universitat Jaume I

Campus Riu Sec, s/n

12071 Castelló

Spain

alcina@trad.uji.es

http://tecnolettra.uji.es

Abstract

Translation technologies constitute an important new field of interdisciplinary study lying midway between computer science and translation. Its development in the professional world will largely depend on its academic progress and the effective introduction of translation technologies in the translators training curriculum. In this paper different approaches to the subject are examined so as to provide us with a basis on which to conduct an internal analysis of the field of Translation technologies and to structure its content. Following criteria based on professional practice and on the idiosyncrasy of the computer tools and resources that play a part in translation activity, we present our definition of Translation technologies and the field classified in five blocks.

Keywords:

Introduction

Translators and teachers of translation have been pioneers in the use of the computer as a tool that is fully integrated into the work process, at least as far as the field of language-related professions is concerned. Although the relationship between translation and computers began with the development of software for machine translation, the real boom of translation technologies started with the development of electronic dictionaries and terminology databases, the arrival of the Internet with its numerous possibilities for research, documentation and communication, and computer-assisted translation tools. The digitization of content generated at the source, and "computerization" of institutions, organizations, private business, professional work places, etc. also played a significant role. The computer has been an integrated part of the infrastructure needed by translators for some time now, but the amount of knowledge and the skills linked to the translation technologies that the translator has to acquire and have a command of is growing every day. Moreover, the steady rise in the number of computers and users, the considerable increase and variety of electronic documents formats, and the rate at which they circulate over the Internet have given rise to a new specialised area called *localisation* – that is, the translation of the documentation, interfaces and help files included in computer software applications and the translation of websites – which requires translators to have a wide, thorough knowledge of computer science that was previously only possessed by specialists.

The translation process is complex and is made up of many sub-processes and tasks of different natures. Numerous computer tools can, or could, be used to enhance the efficiency, speed or quality of some of these tasks or their results. The need to

combine computer processes and tools with those used to translate, together with their continuous development, has given rise to a new field of study known as translation technologies, which has also been called machine translation, computer-aided translation or computer translation. Regardless of the name, we are now at a point where it is becoming essential to draw up a set of principles that help define the limits of the field of study and to select its content. We also need to establish criteria that can help us to classify and structure this new field and this in turn would contribute to its academic development and its spread throughout the professional and educational domains.

In this paper we have gathered different points of view and then taken them as a starting point for the internal analysis and structuring of the field of translation technologies; the different types of classifications are also discussed. Finally we put forward our own proposal for the definition and structuring of the elements that go to make up the translation technologies. Our proposal is based on a selection of components made according to the phases of the translation process and on a classification of the topics that takes into account the type of computer program or resource involved, its function in relation to translation, the amount of knowledge required of the user (expert, basic) and the degree to which the language must be processed. By so doing, our intention is to contribute to the progress and academic development of this new discipline.

Different approaches to structuring the computer applications related to translation

The literature offers several ways of structuring the computer applications related to translation, each of which respond to different criteria or approaches. The

most classical structure is that which divides translation software into two large branches: machine translation (MT) and computer assisted translation (CAT). The steady growth of the field and the rise in the number of tools and resources over the past twenty years have, however, led to the creation of subdivisions within these two branches. Furthermore, the greater attention currently being paid to this field has multiplied the number of points of view from which it can be analysed. It is also to be expected that the divisions existing today will, in the future, expand and become more substantial. Some of the perspectives taken as starting points for the division of the field include the following:

1) According to the degree of automation in the translation process

2) According to the moment in which the tools are used in the translation process

3) According to the level of knowledge about computing required

4) According to their relationship with translation

5) According to the dimension of translation the computer tool is applied to

In the subsections that follow we will outline these different proposals for dividing the field in more detail, since they help us to establish the content of the discipline.

According to the degree of automation in the translation process

In this subsection we draw on the classifications put forward by Hutchins and Somers (1992) and by Melby (1996). Figure 1 reproduces the scale developed by Hutchins and Somers (Hutchins and Somers 1992) to measure the extent to which the

translation process is automated when using the computer tool.

Figure 1. Hutchins and Somers' scale for measuring translation automation (1992)

The term *traditional human translation* is understood to refer to translation without any kind of automation, as it is has been carried out for centuries. At the other end of the scale, *fully automatic high quality translation (FAHQT)* means translation that is performed wholly by the computer without any kind of human involvement and that is of "high quality", which is a goal that was then seen to be a utopia and remains so even today if no controls or restraints are applied to the language in the texts. *Human-aided machine translation (HAMT)* refers to systems in which the translation is essentially carried out by the program itself, but requires aid from humans to resolve specific language problems arising from the source text, or to correct the resulting target text. *Machine-aided human translation (MAHT)* includes any process or degree of automation in the translation process, provided that this mechanical intervention provides some kind of linguistic support. Thus, the machine-aided human translation category includes things like: spelling, grammar and style checkers, dictionaries, thesauri, encyclopaedias and other sources of information that translators can consult either online or on some other computer-readable support. It also contains the systems that make use of these tools and go to constitute a work environment that allows the translator to handle the source and the target text, as well as the storage and retrieval of translations – in other words, what is known as the *translator's workstation*.

Another classification that ties in with this approach is the one put forward by Melby that is based on making a distinction between Level-One and Level-Two tools (Melby 1983, 1996):

➔ Level One. These are tools that the translator can use when the source text is

supplied on paper or by fax. Such tools include voice-to-text systems, word processors, fax machines, e-mail and other telecommunications, word-count and terminology management software. With the aid of these tools, translators can almost eliminate the distance separating them from their clients and thus compete with a local translator. Billing can also be carried out more quickly and accurately, and greater consistency is achieved in the use of specialised terms.

➔ Level Two. These tools require the source text to be in machine-readable form. Such tools include terminology research and automatic equivalence lookup, translation memories, and quality-check tools. All but automatic lookup require an aligned bilingual corpus.

According to the moment in which the tools are used in the translation process

Melby (1998) proposes a classification of the systems that can aid the translator according to the phase of the translation process in which they are applied (before, during or after the translation, or they might be essential throughout the whole process) and according to the level on which they act (terminology level or segment level). These factors allow us to distinguish between eight types of translation technology (Melby 1998):

1. **Infrastructure**. Tools that, despite not being specifically designed for translation, are crucial in the translation environment, such as:

- Document creation/management systems

- Terminology databases

- Telecommunications (intranet/Internet, e-mail, File Transfer Protocol,

web browser, etc.)

2. **Term-level before translation**. Tools that allow the detection of candidates that could be eligible for inclusion in a terminology database. They would not only identify words, as spell checkers do, but also terminological phrases. Terminology research would cover searches on the Internet and in text databases.

3. **Term-level during translation**. Tools that automatically look up terms in the database and show their equivalents in the target language set by the translator. The translator does not need to search for terms in the database manually and can incorporate them into the target text automatically without having to type them out.

4. **Term-level after translation**. Tools that check the consistent use of terminology after a translation has been completed, and tools that flag terms that the translator wishes to avoid using.

5. **Segment-level before translation**. Tools that allow the segments in a source text to be aligned with their matching segments in the target text and indexed.

6. **Segment-level during translation**. This type includes tools for looking up segments in the translation memory and machine translation.

7. **Segment-level after translation**. Tools that detect missing segments (for example, if one of the text segments has not been translated), check grammar and retain the original text format.

8. **Translation workflow and billing management**. These tools do not play a direct part in the translation process itself, but they are extremely important for keeping track of the work done, especially when dealing with large-scale projects. In these cases, it is essential to have tools that help control the variables that can affect the different texts being processed in a project (deadlines, modifications to texts, translation priorities, revision dates, and so on). Figure 2 shows this classification.

Figure 2. Classification of translation tools according to Melby (1998)

According to the level of knowledge required

The interest in instructing Translation and Interpreting students in the use of the technologies has also given rise to a number of more or less explicit proposals for structuring the subject. The LETRAC (*Language Engineering for Translators Curricula*) project[1] put forward a very comprehensive and detailed arrangement. This project was developed with the purpose of defining the technological elements that ought to be included in the European university programs for translators training in order to facilitate their entry on the job market, on the one hand, and to make it possible to reach a high degree of compatibility between the studies in different European universities, on the other. This would clearly provide students with greater mobility and equal opportunities. The starting point for this project was the study of the impact the language engineering have had on translators' work environment.

The structure of translation technologies-related knowledge is divided into three main blocks: 1) Introduction to Computer Science, 2) Information Technologies (IT) and Desktop Publishing (DTP) for Translators, and 3) Language Engineering. Each of these modules is, in turn, subdivided into different topics. The design includes information about the previous knowledge requisites, and it is also stated which modules are essential to translators' instruction and which would be part of their specialised training. Figure 3 shows the diagram representing the interrelations among the modules and their different topics (Badia et al. 1999).

MODULE A: INTRODUCTION TO COMPUTER SCIENCE. This module includes instruction in the physical and logical elements of computers, data storage methods and devices, as well as data management and output. This module is considered to be an

essential prerequisite to acquisition of the knowledge in modules B and C.

MODULE B: INFORMATION TECHNOLOGIES AND DESKTOP PUBLISHING FOR TRANSLATORS.

B1. Desktop Publishing. This includes tools such as Framemaker, Pagemaker or QuarkXpress that enable the translator to produce professional high quality printed material. Physical and logical elements such as scanners and optical character recognition (OCR) applications are also dealt with. Emphasis is placed on textual aspects, graphics and the presentation of printed material. It would require prior knowledge of computing, word processing and the presentation of slides.

B2. Information Technologies and Internet. This area looks at tools that allow information and terminology to be obtained across the Internet, as well as communication tools. Examples include file transfer protocol (FTP), telnet, e-mail, Internet browsers, creation of documents in HTML, networks, and specific Internet services and resources for translators (online dictionaries, databases such as Eurodicautom, machine translation systems, aligned corpora, electronic journals).

B3. Advanced level. This course, which would be optional, would offer advanced training in Web design and desktop publishing, and also expand on general computing topics such as operating systems (advanced DOS/Windows, MacOS, LINUX/UNIX) and programming languages (PROLOG for processing natural language, C, C++, Perl, AWK).

MODULE C: LANGUAGE ENGINEERING

This is considered to be the core module, since it deals specifically with the processing of linguistic data by the computer. It is divided into 6 units:

C1. General aspects and tools. Word processors and checkers, controlled language, project and document management, terminology database management.

C2. Translation-specific aspects and tools. Translation memories and Machine translation systems.

C3. Language formalisation. Structural and functional theories, grammars and parsers, unification-based approaches to grammar.

C4. Machine translation (MT). Historical overview of MT, basic concepts, evaluation of commercial systems, unification-based MT systems, Research Framework Programmes of the European Union about Machine Translation and Natural Language Processing.

C5. Corpus linguistics and low-level analysis. Corpus organisation and design, corpus annotation, low-level parsing, information extraction from corpora and project design.

The LETRAC curriculum was designed to be an exhaustive structured arrangement of all the tools and knowledge that relate computer science to translation. Sequencing the content enables us to distinguish between topics that are considered to be **basic** for the professional performance of translation and **advanced** topics that could constitute a specialisation for students and may be career-oriented (localisation) or research-oriented (natural language processing). It is necessary to notice that the aim of the LETRAC project was to design a well-defined set of LE curriculum elements which can be integrated into a university course for translators, but not the implementation of the course.

Figure 3. Content modules in the LETRAC programme

According to technology's relationship with translation

Within the *Tradumática* group[2], Vilarnau proposed a classification of computer

programs in five groups, depending on how they are related to translation (Vilarnau 2001):

1. **Translation programs**: word processors, assisted translation tools, machine translation tools, desktop publishing applications, HTML editors, software localisation tools, etc.

2. **Translation aid software**: databases, dictionaries and encyclopaedias on optical discs, browsers, spreadsheets, spelling and grammar checkers, etc.

3. **Programs for sending and receiving documents**: e-mail, File Transfer Protocol (FTP), optical character recognition, voice recognition, etc.

4. **Accessory translation software**: compressors, coding programs, chat, web spiders, download management software, etc.

5. **General programs**: antivirus applications, firewalls, operating system maintenance software, etc.

This classification can be represented by five circles as in Figure 4, in which it can be seen that the inner circles are more closely linked to translation than those further away from the centre.

Figure 4. The Tradumática *group's translation technologies circles*

According to the dimension of translation to which the computer tool is applied to

Neunzig (Neunzig 2001:168-194) points out that the application of computer tools to translation involves three dimensions:

➔ Teaching translation

➔ Professional translation

➔   Research into translation

This classification adopts a pragmatic point of view, in that it focuses on what computer programs and resources are used for, and also on the profile of the users of those programs, regardless of the technology they are based on.

Discussion on the different proposals

The proposed classifications suggested by Hutchins and Somers (1992) and by Melby (1983, 1996) have helped to organize the relationships between technology and language processing, and more specifically between technology and translation. The first of the two classifications focuses on the use of technologies as regards the role they play in translation, by establishing different levels on which they help to make the work of the human translator easier and/or replace it. The second, put forward by Melby, adds a classification that also takes into account what the tools are actually like and, more particularly, the sort of data they deal with, thus distinguishing between those that handle data on the term level and those that handle data on the segment level. Moreover, Melby believes that knowledge and use of these tools must be set within the context of the translation infrastructure.

We therefore have a two-level classification of the tools: according to the degree of human involvement (fully automatic, traditional human and computer-assisted translation) and according to how far these tools are involved in translation (required infrastructure, term level and segment level). Yet, the progress made in recent years both in the field of general Information Technologies and in language processing and translation makes it necessary to renew these classifications and to expand their content.

The curriculum that is proposed for training translators in the LETRAC project embraces a large number of tools. Nevertheless, certain topics included in the

curriculum, such as knowledge of programming languages like Prolog or Perl or various operating systems, do not seem to be particularly useful even for today's professional translator. Machine translation also has an excessive relative weight with respect to assisted translation in the curriculum proposed, which is far more widely used by professional translators. Moreover, great importance is given to the theoretical linguistic approach to machine translation (linguistic formalisation theories, machine translation history and systems), when it would be much more interesting, from the professional translator's point of view, to emphasise practical aspects of machine translation such as pre-editing and post-editing of texts. On the other hand, the criteria used to classify tools and resources are not altogether coherent. For instance, dictionaries and terminology databases such as Eurodicautom are dealt with in module B, Information Technology and desktop publishing Technologies, in the Telecommunications/Internet unit, together with tools like ftp (file transfer protocol), telnet, e-mail, Internet browsers, creation of documents in HTML and networks. It would be far more logical, in our opinion, to include instruction in the use of dictionaries and terminology databases in module C, Language engineering, which is described as "tools for processing linguistic data". The same can be said about including training in Prolog or Perl programming languages in module B, since these languages can be of interest in matters concerning Language engineering and processing linguistic data. Their study should remain optional and it should perhaps be stated that they are useful in research and development, rather than in professional translation.

The LETRAC curriculum offers thorough instruction in Language engineering, with a certain extent oriented towards translation. However, the curriculum had an excessive computational linguistics load and it lacked a number of topics that are valuable for the professional translator.

The classification put forward from the *Tradumática* group should be highlighted as being one of the first classifications to focus on professional translation, and on the translator's needs. Thus, it has been claimed that any tool that can be used by the professional translator should be included in this field, and from there they develop their classification in five groups. The proposal (Vilarnau, 2001), however, does not offer any kind of definition or explanation of the criteria used to select the tools that were included or to determine to what extent a given program has to do with translation. It only provides a brief, incomplete list (for example, in all cases the list ended with an "etc.") of computer programs that would go to make up each category. Consequently, the limits between the different groups in the classification are not clearly defined and there are no criteria to decide whether a new tool is to be included or not and, if this is the case, to decide on which group it should be sorted into. The names given to each of the groups ("translation programs", "translation aid software" or "accessory translation software") are not very clear or explanatory either. The fact that the relationship with translation is taken as the reference point for the classification leads to some groupings that would appear surprising if viewed from other fields, such as including word processors in the group called "translation programs". As we see it, regardless of the fact that some tools are acknowledged as being more useful for translators than others, it is difficult classify them according to their "degree of usefulness". It is difficult to say if a translation memory is more useful than a work processor, or if a good electronic dictionary is more useful than a machine tanslation program. At least, it can be subjective or it can depend on the text and another parameters. So, the classification of these tools should not take translation as its *focal* point. Rather it ought to be based on more general criteria that are commonly accepted in closely related fields such as computer science or the language technologies.

Neunzig's classification seems to be a good starting point from which to consider the different areas that translation technologies are applied in, but it is not suitable for classifying the tools, since the same tool can be useful for different purposes (research, teaching and professional translation). What would change would be the use it is put to, the functions utilised by each profile and the methods that would be applied when using them. For example, a text analysis program like WordSmith can be used by a researcher to compare the frequency with which a term is used in different corpora, by a translator to search for contexts that help to determine the meaning of a term, or by a teacher to obtain examples of how a term is employed and to create exercises for his or her students. It will therefore be essential to distinguish between these three dimensions, for example, in order to design courses for one of these users, but such a distinction will not be so useful to classify the content of the field or the tools, since we could well find the same programs in each of the categories.

Having analysed the different classifications and the way the field is structured by a number of authors and at different times with different purposes and then reviewed the strong and weak points in each case, we are now in a position to present our own proposal for the description of the field of translation technologies, how it is structured and the classification of the tools used.

Defining the limits of the field and criteria for selecting its content

On the basis of our review of the literature and our own experience in researching and teaching translation technologies, we describe translation technologies as the field of study that deals with the design and adaptation of strategies, tools and technological resources that help make the translator's job easier, as well as facilitating the research and teaching of such activities. We have also put together a proposal for

structuring the field, and to do so we set out from the following principles: a selection of components suited to the process of translation, structuring the field in a logical manner, and making the selection of those components flexible so as to satisfy different profiles. Here I use the term logical manner in the philosophical sense, and not in the popular sense. In the philosophical sense, logical structuring refers to a way of classification that is based in the similarities of the contents, and not in pragmatic use (like function) or temporal or spacial sequencing.

Selection of materials suited to the translation process

The selection of the components that make up the field of translation technologies must respond to the analysis of the translation process and the determination of the tools that are best suited to perform the tasks involved. In other words, the decision as to whether to **include** certain components or not must be made bearing in mind their application and use from the point of view of the activity and the work process of the professional translator. It must also be based on an up-to-date perspective that takes into account the range of tools currently available to translators.

Thus, the columns in Table 1 contain the activities that go to make up the translation process and the tasks that can be performed, for the same activity, using a computer and suitable tools. We have made a distinction between different phases in the process, although we are fully aware that often in practice no such clear-cut division can be made.

Table 1.- Table comparing the phases of the translation process with and without the help of computer tools

The incorporation of computer tools in the different phases of the translator's work has been examined by different researchers and teachers of translation who have highlighted the different ways such tools could be used in the translator's work environment and have helped to perfect them. Some of the most important studies in this area include those dealing with the translator's workbench (Hutchins 1996; Kugler et al. 1995) or others focusing on more specific aspects of the process such as the use of corpora in both specialised and literary translation (Corpas 2004; Laviosa 2003; Sánchez-Gijón 2005); on the pros and cons of incorporating translation memories into the work of the translator, together with the most appropriate methodology to do so (Benis 1999; Bennett 1994; Somers ed. 2003); the use of e-mail and mailing lists by translators (Alcina 2003); virtual environments (Alcina 2002; Bolaños 2001); or the role played by the Internet in enterprises and the translation process, which is also known as the industry of Teletranslation (O'Hagan 1996; O'Hagan and Ashworth 2002); and the value of the revision functions offered by word processors (Mossop 2001), the pre and post-edition in machine translation (Krings 2003; Allen 2005) to mention just a few of the many fields of interest. Lastly, a number of works have also examined how these tools are to be introduced into the training of translators (Kenny 1999; Folaron 2006). There are also a few books (Austermuhl 2001; L'Homme 1999; Bowker 2002) and edited books (Somers 2003; Dunne 2006; Kugler et al 1995; Gambier et al 2001) that deal with these aspects as a whole. Of course, translation journals includes articles about software applications for translators, and recently also machine translation journals includes articles about the professional aspects of translation (see *Machine Translation News International*, edited by Laurie Gerber).

Classification criteria

Once a criterion has been elaborated for including or excluding elements in the field of translation technologies, the next step we have to deal with is that of ordering all the functions and support tools for translators. To do so, we will take into account whether they have more or less to do with translation itself but without neglecting the **conventions** for classifying tools that have been established and generally accepted by other fields. Thus the structuring is based on logical criteria, that is to say, on the similarities between the actual functions and tools, using criteria such as:

a) The type of data they handle: from letters or numbers up to texts or sets of texts.

b) The type of data processing they perform: formatting texts for printing the final version, carrying out complex calculations, obtaining lists of words or text concordances, performing morphological or syntactic analyses of a text before translating it into another language, term storage, search and lookup, and so on.

c) The function or purpose they are used for: office functions, keeping the accounts in an enterprise, to obtain a preliminary draft of a translation, language consultancy or mediation, and so on.

d) Whether they handle language information or not and, if so, what kind of information: whether the data is labelled, for example, on a morphological, syntactic or lexical level, or if the information refers to one or several languages.

The use of these criteria will bring us to a classification that could be generally accepted and shared with other fields related to translation technology, such as language

technologies and computer science.

Furthermore, we make a distinction between **tools** and **resources**. The word *tool* refers to computer programs that enable translators to carry out a series of functions or tasks with a set of data that they have prepared and, at the same time, allows a particular kind of results to be obtained. Thus, translators use the word processor to write out their translations, their translations can be created and stored with assisted translation software, and terminology database management software can be used to store the terminology that has been collected throughout the translation process.

By *resources* we refer to all sets of data that are organised in a particular manner and which can be looked up or used in the course of some phase of processing. For example, dictionaries or corpora, available either online or on CD ROM, are sets of data that can be accessed in different ways. They normally constitute closed data sets that cannot be expanded by information from the user.

Lastly, in classifying the materials that comprise translation technologies we also have to bear in mind the different levels of **affinity** with translation (i.e. the extent to which they have to do with it). This becomes necessary if we consider the fact that some features are essential for translation but are also crucial to other professional fields (as is the case with e-mail or file decompression programs, for example). Other aspects are closely related to professions linked to language mediation and are shared by other fields such as language technologies (such as electronic dictionaries), and still others are specifically linked to translation (machine translation and computer assisted translation).

Characteristics of the classification

If it is to be of any use, the classification must offer certain characteristics such

as being expandable and flexible. Expandable, on the one hand, means that it must remain open to the innovations that may take place in the future, especially in a field that evolves as quickly as the technologies. Moreover, it has to be flexible, in other words, it must be capable of adapting itself to different user profiles and different situations where the technology is applied (for example, within the context of a particular enterprise or the syllabus of a university subject). It must be borne in mind that, although two tools have been designed to perform the same function, they may well require different degrees of skill from the user. For example, some concordance applications run under the Windows environment while others will only work under Unix or MS-DOS, which are, generally speaking, less familiar to users who are not experts. As pointed out by Abaitua (1999), a professional translator can be placed on varying levels of specialisation as regards use of the technologies[3] and, on the other hand, other profiles can also be distinguished, such as the translatology researcher or the researcher and/or teacher in translatology or in translation technologies, who will have to be in possession of different types of knowledge concerning these technologies.

Furthermore, it is also possible to develop adaptations according to the area of specialisation the translation belongs to – literary, audiovisual, specialised (scientific-technical, legal, administrative) and localisation. Each of these specialised areas requires a thorough command of a certain subset of Translation technology tools and perhaps a very broad knowledge of others.

Although the technologies help to enhance the quality, speed and cost of the results of translation and the process, the fact is that not all translators use the different kinds of technology in the same way. Translation memories, for example, have been very successful and widely accepted, yet these programs will only be useful to translators who work with specialised translation (legal, scientific-technical,

localisation), and not to those who are involved in literary or audiovisual translation. Whether the translator makes use of one or other technology will largely depend on the specialised area he or she works with.

As translation in general becomes more *technified*, translators and interpreters from different specialised topics are beginning to ask themselves how they can take advantage of the technologies and adapt them to their individual needs. As a consequence of this, new programs that fit the specific needs of a particular specialised area are beginning to appear. This is the case of assisted audiovisual translation applications, such as *Subtitul@m*[4], or the translation and terminology tools that now are used also within the field of Interpreting.[5]

This panorama leaves us wondering to what extent and in what way the technologies can be applied in the different specialised areas, and I believe that this will involve another classification that we have to start taking into account with a view to the future, especially in the field of teaching, in order to provide instruction in the use of technologies that are closely linked to the specialised areas of professional translation.

Lastly, the structuring of the field must be carried out in such a way as to allow its adaptation to the translator's curriculum. In developing the curriculum, apart from the structure of the field of knowledge, it is also necessary to take into account other factors such as **students' prior knowledge** and profile, the infrastructure available in the educational centre and the demands of the translation market in a particular setting.

Classification of Translation Technology tools and resources

On the basis of the criteria that we have used to select the materials that will be included as elements of the translation technologies and the criteria that we have outlined for structuring the field from a logical point of view, we now describe our

proposal for classifying the components of translation technologies in five large blocks:

    1.- The translator's computer equipment

    2.- Communication and documentation tools

    3.- Text edition and desktop publishing

    4.- Language tools and resources

    5.- Translation tools

A)   The translator's computer equipment

In this case we are dealing with elements related to the general functioning of the computer, that is to say, its physical components, operating systems (Windows, Linux, McIntosh), computer programs that are concerned with computer maintenance (antivirus, file compression and decompression software), the different data input and output elements and types of information support (floppy drive, CD-ROM and DVD drives), computer peripherals (printer, scanner) and basic utilities software (notepad, simple image processing software, screenshot applications). There are still others that are very useful to translators, such as CD-ROM emulators, which allow the user to *set up* several virtual drives that can be used to introduce a number of dictionaries on CD-ROM.

B)   Communication and documentation tools

These components include the concepts, tools and resources that translators use to interact, through the computer and networks, with their actual or potential clients, with other translators or specialists, or to obtain information and data from other computers or servers. The purpose behind such communications can be to seek translation jobs, settling doubts about a certain job, sending and receiving texts,

documentary research and consulting websites that offer specialised knowledge (bibliographic databases, online encyclopaedias, the websites of universities and research groups). Here, we are dealing with both the use and the adaptation of the information and communication technologies in the translation process. As well as enabling translators to carry out documentary research on the specialised area dealt with in the translation, communication tools will also provide them with good communications with their clients or enable them to offer their services in a suitable form through a well designed webpage. Lastly, these technologies will allow them to communicate with other translators, terminologists or computer experts they must work with on the same project, by using the different Internet services (e-mail, chat, videoconference and Web conference software, networks, file transfer, mail lists, and especially virtual collaborative work environments or private virtual networks). Besides these technologies, the translator will have to know how to use software utilities that enable communication to take place under safe conditions, such as firewalls.

C)   Text edition and desktop publishing

Any tools that are used for writing, correcting and revising texts are part of this block of elements. Especially word processors but also HTML and/or XML editors, validators, Web page design applications. Nowadays, word processors already come with a series of built-in functions like spelling, grammar and style checkers and dictionaries of synonyms in several languages. In addition to these functions, which aid the task of writing a text properly, the processor also incorporates other revision functions that enable other people to make modifications to a document that can later be either accepted or rejected by the author, while keeping the different contributions made by each of the people who have processed the text separate. This function can be used to revise and correct translations. A thorough knowledge of the word processor and its

advanced functions can greatly simplify the performance of repetitive tasks (for example, by using macros) or the job of formatting documents (for example, by creating and utilising templates).[6] In subtitling, for example, each line of text cannot have more than certain number of characters. The translator would have to count the characters in each line by hand, but she/he can now create a template that limits the number of characters per line. Furthermore, it is also a good idea to have a working knowledge of optical character recognition software, which allows a scanned image to be converted into text that can be edited, or applications that can be used to convert files from one format to another, such as pdf to txt or other formats. The texts in machine-readable formats can then be added to translation memories.

These useful functions actually implemented at word processors, maybe will be implemented at the desktop publishing applications and editors included in computer-assisted translation software.

D)   Language tools and resources

This block includes the tools and resources designed for the collection and organisation of linguistic data (texts, vocabulary, terms and phraseological expressions, together with any linguistically interesting information such as the grammatical category or meaning, in the case of words, or the textual genre in the case of texts). It is essentially made up of electronic dictionaries, databases and text corpora.

A distinction must be made between tools and resources. Tools allow translators to introduce and manage their own linguistic information, whereas resources are sets of previously gathered linguistic data which are made available in some electronic format so that they can be used or looked up by translators. Thus, Multiterm and WordSmith are tools that help translators in terminology database management or text analysis. The

British National Corpus, Eurodicautom or Termium, however, are language resources that the translator can use to look up text corpora, in the first case, and terminology in the other two. (See the distinction between tools and resources at the Classification criteria above).

E)   Translation tools

These are tools that play a part in the actual translation process, that is, that are specifically designed to work with at least one source text and one target text at the same time and establish relationships between both texts on a segment level or whole text level. In some cases they are combined with another type of software that belongs to other categories, such as word processors or terminology databases. Everything seems to indicate that in the future these different programs (translation, word processor, databases) will become far more integrated, as has happened with word processors, which now include spelling and grammar check functions. In any case, the point to be highlighted as regards these applications is that they are capable of managing a source text and its corresponding target text at the same time.

This category embraces assisted translation programs (which include translation memory management software, terminology databases and word processor) and machine translation programs.

Translation programs are expected to diversify in the coming years in order to meet the needs of the different areas of specialisation encountered in translation. Computer-assisted translation (essentially consisting of translation memories) and machine translation are useful in specialised translation, that is, when dealing with texts containing a large amount of terminology and repeated structures. Computer-assisted translation tools are also of great importance for localisation purposes, in other words, for translating computer programs, video games and websites. In addition, other

programs that aid audiovisual translation or interpreting have also recently begun to appear and it is to be expected that as time goes by more and more programs will be marketed that are aimed at facilitating specific tasks in different specialised areas of translation. To mention just a few examples, literary translation could benefit from having access to corpora of literary texts, and legal translation would be aided by having access to databases containing standardised models of documents used in the legal area.

Conclusions

The translation technologies constitute a new field that requires theoretical consideration and discussion to achieve a structure of its own and internal coherence, which will be an essential step on the way towards its academic and professional consolidation and development. In this paper we have attempted to contribute to reaching these goals by conducting a selection of components based on professional practice and developing a structure that is grounded in criteria that allow for a clear division by topics and which, at the same time, can be adapted to different professional and academic situations.

The rapid development of the translation technologies and the array of tools and resources translators have available to them means that the classification we have proposed here will have to be quickly expanded and subdivided too, and we will have to keep on the watch to capture the new distinctions that will be have to be made. Moreover, we can also foresee seemingly contradictory developments: on the one hand, the development of translation tools and resources that integrate different functions that translators nowadays have perform using different tools and, on the other hand, the creation and development of translation tools and resources that are restricted or specific to a particular specialised area of translation, such as audiovisual, legal or

literary translation.

Alcina, A., in press, Translation Technologies: Scope, tools and Resources, Target: International Journal on Translation Studies.

Notes

[1] LETRAC research project was developed by a set of researchers from different European Universities (1998-1999). The project was funded by European Union and the results can be viewed at < http://www.iai.uni-sb.de/iaien/en/letrac.htm >.

[2] Tradumatica is a research group in Barcelona on Information Technologies applied to translation. See <http://www.fti.uab.es/tradumatica/grup/>

[3] Abaitua (1999) distinguishes between six levels of competence in translation technologies: consultant, user, instructor, evaluator, manager, developer. Also, he says that anyone with a university degree in Translation should be qualified to carry out at least the first three.

[4] Subtitul@m is a subtitling software for translation students created by Universitat Autònoma de Barcelona.

[5] See Jekat and Tessiore (2000) for machine interpretation systems, or the edited book by Manuel (ed.) 2003.

[6] Note the significant work done by Jost Zetzsche to train and educate translation professionals in this domain.

Works cited

Abaitua, Joseba. 1999. Is it worth learning translation technology? Paper presented at *3rd Forum on Translation in Vic. Training Translators and Interpreters: New Directions for the Millennium. May 12-15, 1999*, Universitat de Vic, Spain.

Alcina, Amparo. 2002. "Tutorizar trabajos terminológicos en un entorno virtual de trabajo colaborativo". *Sendebar* 13. 169-181.

Alcina, Amparo. 2003. "Encouraging the use of e-mail and mailing lists among translation students". *Meta* 48:4. 634-641.

Allen, Jeff. 2003. "Post-editing". Harold Somers, ed. *Computers and Translation. A Handbook for Translators*. Amsterdam: John Benjamins

Austermühl, Frank. 2001. *Electronic tools for translators*. Manchester: St. Jerome Publishing.

Badia, Toni, Karl Heinz Freigang, Johann Haller, Christoph Horschmann, Dieter Huber, Belinda Maia, Ursula Reuther and Paul Schmidt. 1999. LETRAC Curriculum Modules. < http://www.iai.uni-sb.de/iaien/en/letrac.htm >

Benis, M. 1999. "Translation Memory". *Bulletin of the Institute of Translation and Interpreting* April - May.

Bennett, Paul. 1994. "The Translation Unit in Human and Machine". *Babel* 40:1. 12-20.

Bolaños, Alicia. 2001. "Las unidades de traducción técnica en red: una experiencia formativa". Carmen Valero Garcés and Isabel de la Cruz Cabanillas, eds. *Traducción y nuevas tecnologías. Herramientas auxiliares del traductor*. Alcalá de Henares: Universidad de Alcalá. 153-162.

Bowker, Lynne. 2002. *Computer-Aided Translation Technology. A Practical*

Introduction. Ottawa: Presses de l'Université d'Ottawa.

Corpas, Gloria. 2004. "La traducción de textos médicos especializados a través de recursos electrónicos y corpus virtuales". Luis González and Pollux Hernúñez, eds. *Las palabras del traductor*. 137-164.

Dunne, Keiran J. ed. 2006. *Perspectives on Localization*. Amsterdam: John Benjamins Publishing.

Gambier, Yves and Henrik Gottlieb. eds. 2001. *(Multi) Media Translation*. Amsterdam: John Benjamins.

Hutchins, W. John. 1996. "Computer-based translation systems and tools". *The ELRA Newsletter* 1:4. 6-9.

Hutchins, W. J. and H. L. Somers. 1992. *An Introduction to Machine Translation*. London: Academic Press. <http://ourworld.compuserve.com/homepages/WJHutchins/IntroMT-TOC.htm>

Jekat, Susanne J. and Lorenzo Tessiore. 2000. *End-to-End Evaluation of Machine Interpretation Systems: A Graphical Evaluation Tool*. Hamburg: University Research Center on Multilingualism.

Kenny, D. 1999. "CAT Tools in an Academic Environment: What Are They Good For?" *Target* 11:1. 65-82.

Krings, Hans P. 2001. *Repairing texts : empirical investigations of machine translation post-editing processes*. Kent, Ohio: The Kent State University Press.

Kugler, M., K. Ahmad and G. Thurmair eds. 1995. *Translator's Workbench. Tools and Terminology for Translation and Text Processing*. Berlin: Springer].

Laviosa, Sara. 2003. "Corpora and the translator". Harold Somers, ed. *Computers and Translation. A Handbook for Translators*: John Benjamins

L'Homme, Marie Claude. 1999. *Initiation à la traductique*. Brossard (Québec):

Linguatech.

Manuel, Jesús de. ed. 2003. *Nuevas tecnologías y formación de intérpretes*. Granada: Atrio.

Melby, Alan K. 1983. "The Translation Profession and the Computer". *CALICO Journal* 1:1. 55-57.

Melby, Alan K. 1996. "Machine Translation and Other Translation Technologies". *Annual Review of Applied Linguistics* 16:0267-1905. 86-98.

Melby, Alan K. 1998. Eight Types of Translation Technology. Paper presented at American Translators Association ATA 39th Annual Conference. November 4-9, 1998, Hilton Head Island, South Carolina. Available at: <http://www.ttt.org/technology/8types.pdf >.

Mossop, Brian. 2001. *Revising and Editing for Translators*: Translation Practices Explained. Manchester: St. Jerome.

Neunzig, Wilhelm. 2001. La intervención pedagógica en la enseñanza de la traducción on-line – cuestiones de método y estudio empírico, Departament de Traducció i d'Interpretació, Universitat Autònoma de Barcelona: PhD Thesis.

O'Hagan, Minako. 1996. *The Coming Industry of Teletranslation*.vol. 4: Topics in Translation. Clevedon: Multilingual Matters.

O'Hagan, Minako and David Ashworth. 2002. *Translation-Mediated Communication in a Digital World*. Clevedon: Multilingual Matters.

Sánchez-Gijón, Pilar. 2005. *L'ús de corpus en la traducció especialitzada: Compilació de corpus ad hoc o extracció de recursos terminològics*. Gerona: Documenta Universitaria.

Somers, Harold ed. 2003. *Computers and Translation. A Handbook for Translators*: John Benjamins. *Translation Studies*.

Vilarnau, Joan. 2001. La Tradumàtica. Delimitació conceptual i importància en el procés de formació del traductor, Universitat Autònoma de Barcelona: Research paper.

Zetzsche, Jost. 2005. *The Translator's Tool Box. A Computer Primer for Translators*. Winchester Bay: International Writers' Group.

Interesting Webs

Asia-Pacific Association for Machine Translation (AAMT). <http://www.aamt.info>.

Association for Machine Translation in the Americas (AMTA).

<http://www.amtaweb.org>.

eCOLORE. <http://ecolore.leeds.ac.uk>.

eCOLOTRAIN. < http://ecolotrain.uni-saarland.de/>.

European Association for Machine Translation (EAMT). <http://www.eamt.org>.

European Language Resources (ELRA). <http://www.elra.info>.

Localisation Research Centre (LRC). < http://lrc.csis.ul.ie>.

Localization Industry Standards Association (LISA). < http://www.lisa.org>.

Multilingual Computing and Technology. <http://www.multilingual.com>

Observatorio de Tecnologías de la Traducción (OTT). <http://www.uem.es/web/ott>