[*Translating and the Computer 28*, November 2006 (London: Aslib, 2006)]

Translation Memories Survey 2006: Users' perceptions around TM use

Elina Lagoudaki
Humanities Programme
Imperial College London
United Kingdom

e.lagoudaki@imperial.ac.uk

Abstract

Translation Memory (TM) systems have been under the spotlight of translation technology research led by both software developers and academic institutions. Both ends try to find ways to maximize the benefits deriving from the use of these tools, whether those translate into productivity enhancements or cost savings. The involvement of the user in these efforts has always been problematic. It is usually too costly, it delays the development of the product because it takes time, and it requires a well designed mechanism to be in place that facilitates the communication between the user and the developer. Naturally, many developers cannot afford to set up such capability, thus they risk producing TM tools that fail to correspond to the needs of translation professionals.

The Translation Memories Survey 2006 (abbr. TM Survey 2006), reported in this paper, was initiated with a view to acting as this very channel of information deriving from users (or potential users) of TM systems. The main purpose behind it is to present the users perspective about TM systems and to supply data on the application domain, that is, information on the procedural aspects of the translation activity, on frequent work practices and on the tasks related to TM systems. It reports on the factors that affect TM use and offers an evaluation of the most commonly used systems according to functional and non-functional criteria. The results also reveal a range of future directions in TM research as those are envisioned by translation professionals.

1. Introduction

Fifteen years ago the first commercial Translation Memory tools were made available to translation professionals with the aim of improving their working life. Since then, TM tools have maintained their objective of assisting and accelerating the translation process (and many of its surrounding activities), but the technology employed has passed through various levels of sophistication in a constant quest for optimal solutions to match the translation industry's demands.

As the years go by, the technology is not the only thing which has been evolving. The needs and expectations of the users of TM systems have been evolving too, either because the information

economy has imposed new rules to their work environment, which brought changes to their work practices, or because they are becoming increasingly more competent in computer usage, hence more demanding in relation to computer applications on offer. Both technology and user expectations drive developments and innovation in TM systems. But whereas technology is fairly easy to manipulate and apply to TM systems, capturing (or even understanding) the needs of translation professionals to whom the systems are addressed proves to be a very difficult task, often undervalued by TM developers. Hence, it is no surprise to see that many of the existing commercial TM systems are technology-driven applications (e.g. with an abundance of useless features and a complex, impractical and difficult to learn user interface), rather than user-driven applications.

According to data deriving from interviews with several TM developers, only a few spend adequate time on extensive consultation with users during the design phase of a TM system (which constitutes the initial stage of TM systems development). For the majority, the only opportunity they offer to the users to express their views on a TM software before it comes onto the market is during the last stage of development, the testing phase, when users (e.g. translators, terminologists, project managers, company owners, etc.) are invited to provide feedback on an almost finished product with limited possibilities for changes. This kind of feedback is, obviously, restricted to only fixing bugs in the system. Hardly ever does a TM software developer start redesigning their system from scratch after a serious problem has been discovered in the architecture during the testing phase. The response to such an event is usually quick workarounds and postponing the solution to the problem until the next version release.

In some rare but existing cases a TM software package comes onto the market without any prior involvement of translators in the development. All the testing is reserved for the software engineers who perform the standard quality checks that a system has to pass before it enters the market. Normally, these vendors invite feedback from the users after the release. But, clearly, this feedback comes too late for improving a system that many translators have already bought. The rest of the translators who are interested in this system will have to wait until the release of the updated version, which according to an approximate estimate is a year on average.

Capturing the voice of the user while designing or improving a TM system is ideally an ongoing effort, that requires having in place an effective mechanism of communication between the developer and the user (Kotonya, 1998). There are a variety of techniques that each TM developer can use to engage the users in their system development process (depending on the resources they can spend), such as: focus group sessions, online fora that can give regular feedback on a product, interviews, surveys, testing the software by in-house teams of users before it reaches the market, and so on. Unfortunately, not all TM developers are able to build a strong channel of input deriving from the users of their systems. But making assumptions about the users' desiderata can only increase the risk of a misalignment between the needs (which designate the intended use) and the functionality of the system.

2. The state of the art in Translation Memory systems

Since the appearance of the first Translation Memory systems in the market, thousands of translation professionals have embraced the TM technology after realising the benefits it can bring in terms of productivity, cost savings and quality of the translation output for all the players in the translation industry: translators, language service providers and their clients. The success story of TM systems continues today by the increasing dissemination of this particular technology and the perception of TM systems as a highly profitable product. These days, more than ever before, a great number and variety of TM tools is available to translation professionals, with new tools entering the market every few months[1], intensifying the competition in the TM tools arena.

Naturally, every TM system has some unique features that differentiate it from the rest and performs certain tasks better than other systems; that is why it is preferred by a certain share of the consumer market. However, all TM systems share a common purpose of use (i.e. the deployment of existing translation resources in a new translation project) and they are all based on the concept of the translation memory[2] which is the core element in every TM system. Created on the basis of this concept, the system stores and indexes previously translated content in an organised way, so that it can later retrieve from it as much information as possible (i.e. a match for a current translation unit) when queried by the user.

The major differences between TM systems lie in the methods used in the key internal processes of a TM tool. These processes govern text segmentation, alignment, indexing, search and retrieval.

Most TM tools perform the segmentation and alignment processes before indexing the text. They segment the source and target text into translation units (those can be entire paragraphs, sentences or phrases), they align them and then they store and index the pairs of translation units in the TM database. However, some tools (such as MultiTrans and LogiTrans) follow a different approach, called the full-text approach. Instead of segmenting the texts at the beginning, they store them as full bitexts and index them in the TM database using the character-string-in-bitext (CSB) technique (Gow, 2004). Once the bitexts are in the database, they are aligned at paragraph level. This approach has two main advantages compared to the traditional method: a) the faster creation of a large TM database containing previously translated material and b) the retention of context for any match found and suggested to the user (Gervais, 2002).

There are also tools that do not require the use of a TM database for the storage of the previously translated material. STAR Transit, for example, instead of creating an external TM database, builds and makes use of a 'virtual' translation memory by associating the translated files that reside in any directory on one's computer. The user chooses the translated documents that are relevant to his/her

[1] Known new arrivals in 2006 are: Cafetran 2006 by Igor Kmitowski, MemoQ by Kilgray and Lingotek by Lingotek, Inc.

[2] The concept of translation memory as described here should not be confused with the translation memory database. Some TM systems do not have a TM database at all, while others have an index instead.

project, and the system, after extracting the text into XML files, creates the associations between them. The main advantage of this idea is the absence of issues related to database maintenance and to the security of data residing in a database.

The most important difference between TM tools is in the match retrieval techniques. These techniques determine how efficient the TM tool is in retrieving all available exact or fuzzy matches for a source segment (match recall) and how efficient it is in retrieving the correct exact or fuzzy matches for the source segment (match precision). Some of the latest match retrieval techniques employed are: a) character-string-based matching and b) linguistically enhanced matching. A number of TM tools which implement the first technique (such as DéjàVu), look for matches in the sequence of character strings of each segment and can recognise matches not only at segment level but also in sub-parts of the segments that exist in the TM database. DéjàVu extends this method by also implementing Example-Based Machine Translation techniques to improve its fuzzy matching. For example, if, for a source segment, it finds two sub-segments that exist in two different segments stored in the TM database, it puts together the two sub-segments to form a new segment to suggest as a match.

Implementing the second matching technique requires the introduction of linguistic information to the system. SIMILIS, for example, after segmenting the source and target texts at sentence level, runs a linguistic analysis and further splits each sentence into syntactic units ('chunks') attaching grammatical annotations to them (this is performed with the help of monolingual lexicons and algorithms that can recognise grammatical categories) (Planas, 2005). It then indexes those as translation units as well. So, every time the system searches for a match, it looks not only at the sentences, but also at the chunks (thus increasing the possibilities of finding one), and especially those chunks that are in the same grammatical category as the source segment (thus increasing the possibility of finding the right one). Masterin, on the other hand, segments the source and target texts in a flexible way according to the examples available in the TM database ('Knowledge Base') provided. Each segment is annotated with grammatical information and constitutes a grammatical 'translation pattern'. So, matches are sought by a deep-structure pattern recognition method which looks beyond the surface appearance of segments. If several matches are found, the system determines the best match by using semantics (with the help of a built-in lexicon) and/or examining their use frequency or domain information. In the case where no match is found, the system constructs and suggests a fuzzy match from the available resources in the database by applying translation heuristics (Grönroos, 2005). Although the second match retrieval technique is reported to produce improved results in terms of both match recall and precision, it has a significant disadvantage over the first method. TM systems that use linguistically enhanced matching techniques are language dependent since they rely on the built-in language resources. Consequently, such systems can work only for a small number of language combinations for which adequate language resources have been developed and incorporated to the system.

Another major difference between TM tools has to do with the translation environment they allow users to process their translation in (Zetzsche, 2005:164-7). Some TM tools work as add-ins (through macros) to Microsoft Word, while others provide their own translation processing environment, usually in a

tabular way. TM tools that belong to the fist category are: Wordfast, MultiTrans, Logoport, Metatexis, Trados and Fusion. DéjàVu, Heartsome, MemoQ, STAR Transit, SDLX and across are some of the systems that belong to the second category. Both these approaches to design have respective advantages and the choice of one design over the other is usually a matter of the user's preference.

Finally, it is worth noticing the interest in developing open source TM software. Some of the efforts have led to the release of working products (such as Transolution, ForeignDesk, OOxlate), and some others have been less fruitful. Among the open source TM systems, the most successful ones that have been under continuous development up to the present day are: Omega-T and the Open Language Tools.

All the above TM tools are different in design and technology employed. However, their development always departs from the same point of reference: addressing the needs of the modern translation professional. And because the identification and understanding of these needs plays a significant role in choosing both the right design and technology (either when developing a new TM system or improving an existing one), it is essential to have access to a pool of relevant information about the end users of the system and their work practices. TM Survey 2006 serves precisely this purpose: to enrich the knowledge that TM developers and other researchers have about translation professionals and the aspects of their work in need of TM technology solutions.

3. TM Survey 2006: Design methodology and goals

As in every information elicitation process which precedes the collection of requirements for a system (Griffin et al., 1993:1-27), there were several important issues that had to be considered carefully during the design of the questionnaire. For instance: how can one get information on what the users need, if they do not know what they need? And if they do know, how clearly can they express themselves? Also how can the researcher be sure that he/she interprets what the respondent says in the correct way? Or vice versa, how can the researcher be sure that the respondent understands the questions asked the same way as he/she does? In order to minimise any misunderstanding or loss of information due to the issues above, a pilot study was carried out prior to the survey being designed.

The pilot study included interviews (10 in total) and a focus group session (8 participants in total) carried out at Imperial College London with a mixture of translation professionals. In both activities, the questions initially administered by the researcher were tested and amended where necessary, according to the feedback they received from the interviewees. This was done in order to produce simpler, clearer, less ambiguous questions that the sample unit of the survey would later have no difficulty in responding to.

Other ways to mitigate the above risks were: a) the use of as little technical jargon as possible in the questions, b) the standardisation and ordering of questions in such a way that a question does not influence the response to subsequent questions and c) the suggestion of options (formed in collaboration with the participants of the pilot study) as well as an open space for a different option as

an answer to some questions. The latter technique is proven to facilitate the flow of ideas for people who cannot think of any answer when they first read a question (Sommerville, 2004: 148-53). An effort was also made to avoid lengthy and irrelevant (to some groups of professionals) questions, and to this end a user-friendly tool[3] that regulated dynamically the appearance of the questions was used. The average time needed to complete the survey was only 7 minutes.

The TM Survey 2006 was launched online on 1st July 2006 and closed on 1st September 2006. It was actively promoted to user groups' fora, to translation and localisation companies individually, as well as their associations, to translators' associations, to academic institutions that train translation professionals and to organisations and public authorities that have translation departments. The survey was available only in English, which means that it is biased towards English-speaking translation professionals. To obtain the desired number of responses (min. 500) in the shortest time possible, the Internet was chosen as the medium of carrying out the survey. This means that the survey is also biased towards those who had access to the Internet and could respond to the online survey.

The survey was addressed to all types of translation professionals (such as translators, terminologists, project managers, reviewers, subtitlers, etc.), as it was intended to examine inter alia what is the penetration of TM technology for all sub-groups of professionals and what are each one's attitudes towards TM systems.

The TM Survey 2006 succeeds the research carried out by other independent bodies which have conducted surveys on users' perceptions around TM systems. Some of the more recent known surveys have been: the LISA (Localisation Industry Standards Association) 2002 Translation Memory Survey[4], the eColore Translation Memory Survey 2003[5] and the LISA 2004 Translation Memory Survey[6]. The two surveys carried out by LISA attempted to distil information on similar research questions as the TM Survey 2006, but their main difference is that they were addressed to corporate users of TM systems (companies, translation departments within organisations), hence the focus of their results was the implementation and usage of TM technology in relation to cost savings and gains in quality and faster delivery of services. The TM Survey 2006, on the other hand, covers a wider and more varied sample unit, with the majority of its respondents being freelance translators. Therefore, the focus of this survey is the use of TM technology in relation to the utility of TM systems as well as the productivity and quality gains they may bring to the work of the user. The eColore survey covered a wide range of TM users and had a similar focus to the TM Survey 2006, but instead of providing an appraisal of different TM tools, it was more interested in investigating training issues deriving from the use of TM systems.

---

[3] The tool that was used to create and publish the TM Survey 2006 was SurveyMonkey (www.surveymonkey.com)

[4] Information on the LISA 2002 Translation Memory Survey can be found on USA's website (URL: http://www.lisa.org/products/surveys/tm02survey.html)

[5] Information on the eColore Translation Memory Survey 2003 can be found at http://ecolore.leeds.ac.uk/xml/project/news.xml?lang=en

[6] Information on the LISA 2004 Translation Memory Survey can be found on LISA'S website (URL: http://www.lisa.org/products/surveys/tm04survey.html)

The main goal of this survey is to shed light on the user's perspective about TM systems. More specifically, the aims of the TM Survey 2006 are:

- to establish the needs of translation professionals via their practices and working habits during the translation process;
- to reveal the tasks related to TM use;
- to distinguish the profiles of different TM user groups according to criteria such as the type of tasks they perform, their professional status, their years of working experience, their computer usage competence, etc.
- to provide an insight into the work environment in which translation professionals carry out the translation activities today;
- to estimate the TM technology penetration in the translation market;
- to help understand the reasons behind low usage of TM technology and to discover missed opportunities for reaching potential users;
- to uncover user satisfaction levels for existing TM systems;
- to open the way to new ideas about future systems and identify possibilities for expanding the functionality and scope of use of TM systems.

4. Survey results and analysis

874 translation professionals from 54 countries responded to the TM Survey 2006 during the two-month period. The analysis of results was accomplished using statistical methods and it is presented below[7] in six separate sections. The first section offers a categorisation of participants into groups depending on their occupation, working status, years of experience, age, qualifications and computer usage competence. This helps in drawing relationships between different types of users and establishing user profiles. The second section supplies data on the specialisation of users, their annual volume of translation work, the file formats most frequently used, the frequency and reasons for using the Internet and the translation research methodologies used. The third section reveals the percentage of translation professionals using a TM tool and the reasons for doing so (as well as what prompted the use). The usage rate is then correlated with the data of the previous sections to generate conclusions on which are the types of users who use TM tools most and for which type of content. It also reveals and analyses the reasons for those who are not using a TM tool, reports the trend to try out or buy a TM tool, and discusses the limitations of TM tools that cause low TM usage. Section four shows how TM systems are used by translation professionals, what are the perceived benefits and how length of TM usage and volume of work influence the level of TM usage. In the fifth section, users reveal their TM tool preferences and evaluate the tool they use most, according to a variety of criteria, such as functionality, usability, efficiency, learnability, value for money and customer support. The sixth and last section offers an insight for the future. Users rate the importance of some possible features for future TM systems and give their opinion and ideas about future developments.

[7] It must be noted that because of the large amount and complexity of results and due to the length restrictions of this report, only the key findings are presented and analysed in this paper.

4.1 Participants' profiles

The survey attracted the interest of translators more than any other type of translation professional, perhaps because TM systems are addressed primarily to them. The respondents who declared themselves to be translators totalled 785, whereas 85 were project managers, 59 reviewers or QA managers, 33 terminologists, 28 subtitlers, 13 interpreters and 64 other various translation professionals, such as DTP specialists, graphic designers, web authors, etc. Respondents could choose more than one profession, and in fact 16% indicated that they performed more than one role. This could support the notion that there is not yet a clear-cut distinction between the different types of professionals, and because of this, an individual with an extensive set of skills may be engaged in any of the activities within the translation process.

Regarding their working status, the greatest number (48%) were freelancers working independently without an agency, 19% were freelancers working closely with an agency, 6% freelancers working cooperatively with other freelancers, 8% company owners, 9% company employees in translation/localisation companies and 10% company employees in companies/organisations of other sectors.

From the correlation of the above data, it turns out that subtitlers and interpreters are more likely to be freelancers, whereas project managers are normally employees. Surprisingly, a minority of project managers are also freelancers (see Figure 1).

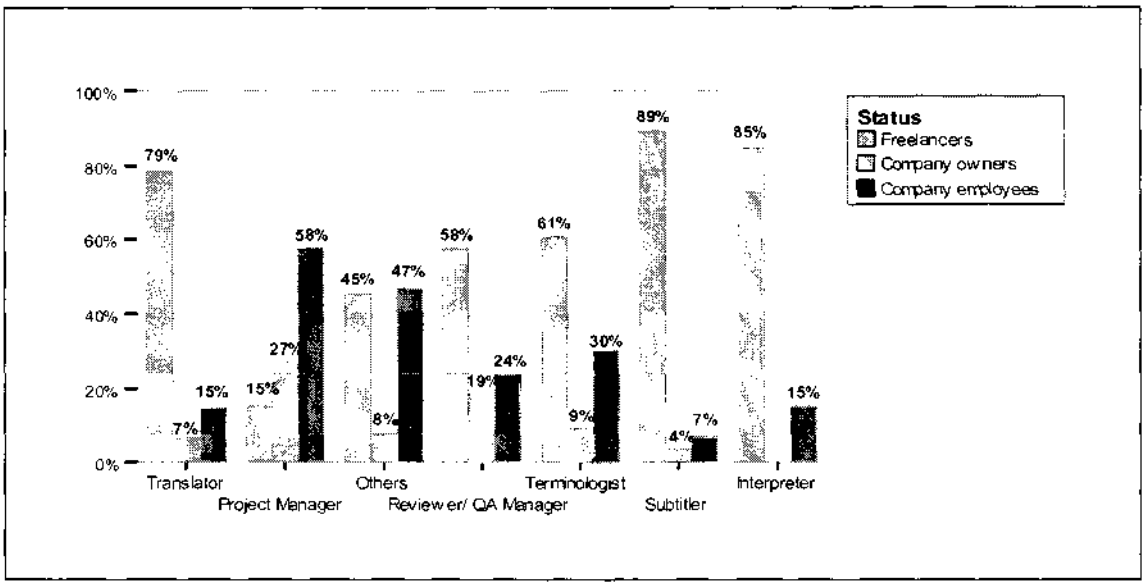

Figure 1: Professions in relation to working status

Asked about their qualifications, only 11% reported not having a professional qualification relevant to their job. In terms of those who did have a relevant qualification, 68% had a university degree, 13% had a certificate by a professionals' association, 3% a certificate by a private institution and 5% had another type of qualification (see Figure 2). These results seem very positive for the translation industry, as they provide evidence to the fact that translation activities are performed by qualified professionals, rather than people with purely practical language skills.

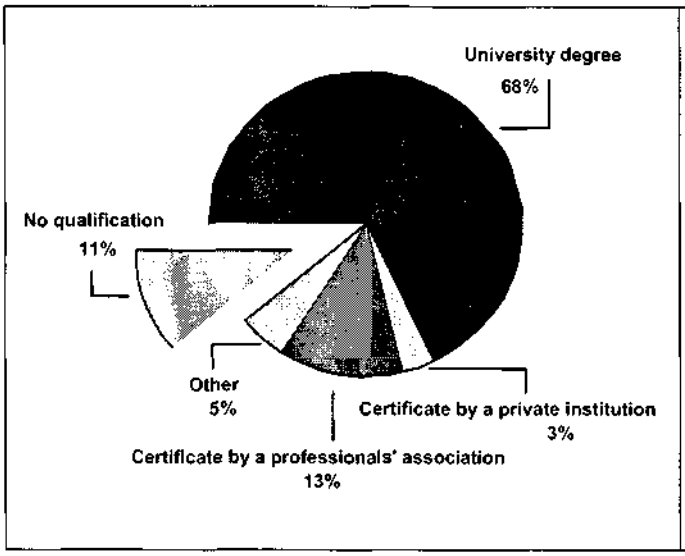

Figure 2: Professional qualifications

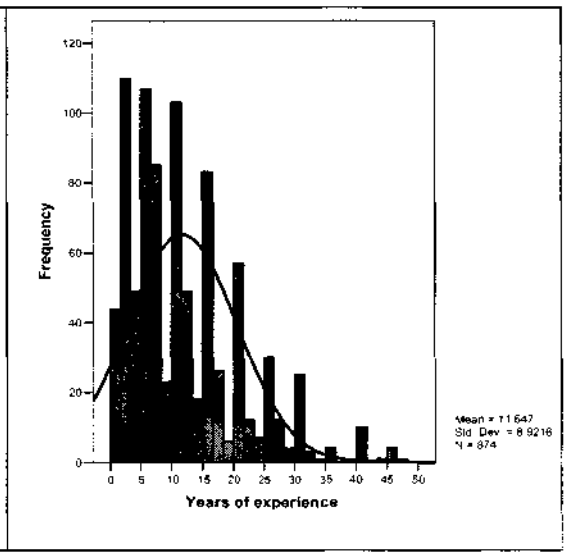

Figure 3: Years of work experience

In terms of years of work experience, respondents reported having from 1 to 50 years of experience, with an average of 12 years (see Figure 3). 29% had less than 5 years of experience, 28% had 5-10 years, 17% 10-15 years, 12% 15-20 years, 10% 20-30, 2% 30-40 and 1% 40-50 years.

Regarding their age, 29% of respondents fell in the 31-40 age group, 21% in the 20-30, 24% in the 41-50 and 25% were over 50 years old.

When they were asked to rate their general computer usage competence, the majority of respondents (64%) reported a 'good' level of computer skills, whereas 30% reported an 'excellent' level and only 6% reported that they had an 'adequate' level of computer skills. This shows that translation professionals have reached a certain maturity in using computers, therefore one should expect that they should feel more confident in applying TM tools to their work.

4.2 Work environment and practices

In order to distinguish further the profiles of the users in relation to the characteristics of their work, they were asked to identify their area of specialisation and the file formats that they work with. The majority of respondents (61%) indicated the technical (including medical and scientific) content as the main area of specialisation. 9% specialise in legal content, 8% in marketing material, 4% in financial, 3% in literature and 15% indicated other specialisation on subjects like history, social sciences, business and administration (see Figure 4).

The most commonly used file formats were, as expected, Word, Excel, PowerPoint and text files which accounted for 96% of the file formats reported in the survey. Respondents had the opportunity to choose more than one file format. PDF files came in second place accounting for 43% of file formats chosen. XML and HTML files accounted for 26%, hardcopy documents for 19%, FrameMaker files for 9%. DTP files for 3%, and various other file formats such as Java properties files, Windows resource files. audiovisual files and TM proprietary file formats accounted for 7% (see Figure 5).

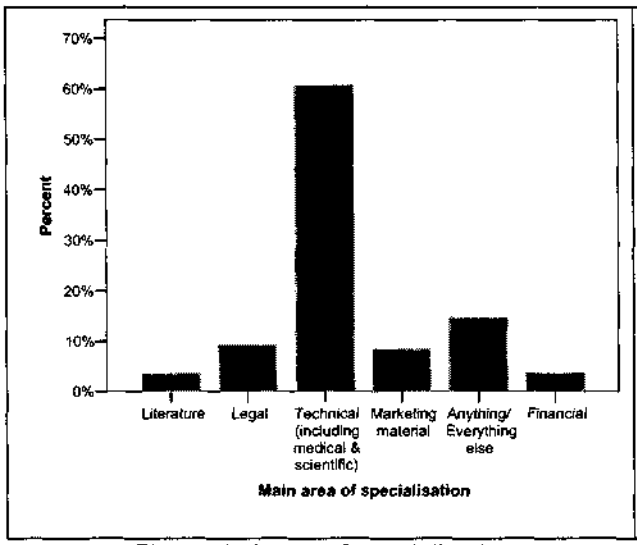

Figure 4: Areas of specialisation

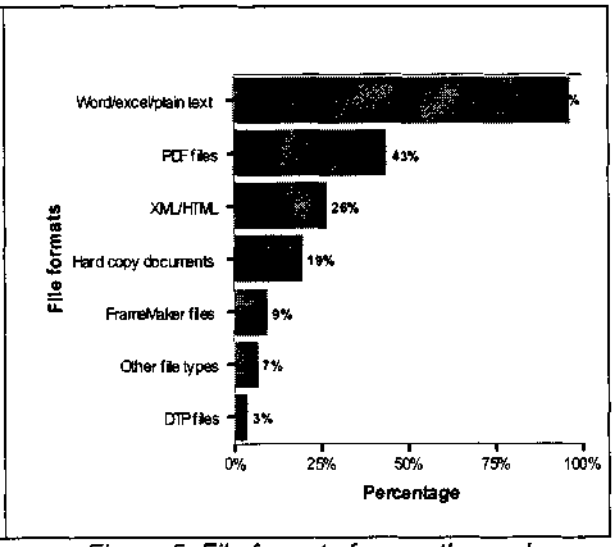

Figure 5: File formats frequently used

Individuals were also asked about the operating systems that they use. As expected due to the dominance of Windows in the OS market and to the fact that most TM systems run on the Windows platform, the vast majority of respondents (92%) use Windows, while Mac users account for 4%, Linux users for 2% and another 2% reported using a variety of operating systems, including any of the above and others such as FreeBDS, BEOS and CP/M.

In terms of Internet connection, the entire sample of respondents indicated that they had access to the Internet. Most of the respondents (58%) have a DSL connection, 19% have a cable connection, 7% have a T1 or faster connection, 5% have ISDN and 5% have a 56 Kbps modem connection. These results show that the majority of translation professionals are moving towards faster connections to the Internet, as they recognise its importance as a valuable tool for their work.

Uses of the Internet	Percentage
communication (reception, processing and delivery of translation orders, communication with colleagues/clients)	87%
source of knowledge (information on specialised areas, reference texts)	81%
source of linguistic information (terminology, usage examples, contexts)	77%
to download other software	31%
to download translation sofware	29%
for other use	10%

Table 1: Common uses of the Internet by translation professionals in relation to their work

In fact when they were asked what are the most common uses of the Internet in relation to their work, they specified a variety of uses (see Table 1). Communication is the primary one (87%), followed by the use of the Internet as a source of knowledge (81%) and as a source of linguistic information (77%). 31% of respondents use the Internet to download software, whereas 29% uses it to download translation software in particular. 10% cited using the Internet for other tasks, such as to perform online translations, to maintain a personal website and to promote their services.

In order to allow us to delve deeper into the work practices of translators, individuals were asked to specify the order of research methods they employ when they are faced with the scenario of not knowing the translation for the source sentence (or phrase or word) they have (and after not being able to find it in their translation memory, if they use a TM tool). 30% of respondents reported looking first in the dictionaries that they have in CD-ROMs (see Figure 6). This is perhaps explained by the fact that translators may not have immediate access to the Internet from their work location so that they can search in online resources, or they might find information on CD-ROMs easier to access and search through, more than any other resource (electronic or hardcopy). 21% resort to the Internet, searching in online glossaries or dictionaries, 17% look in hardcopy dictionaries, 15% use Internet search engines (such as Google) to find the translation, and 13% search in old translations or glossaries. Some of the less followed methods are: searching in monolingual or bilingual texts on the Internet, asking a colleague or a subject expert and searching in printed documents on the subject. The later methods are less frequently used because perhaps the fist methods are usually very effective in providing solutions to translation problems, therefore the latter methods are normally used as last resorts if none of the previous has been fruitful.

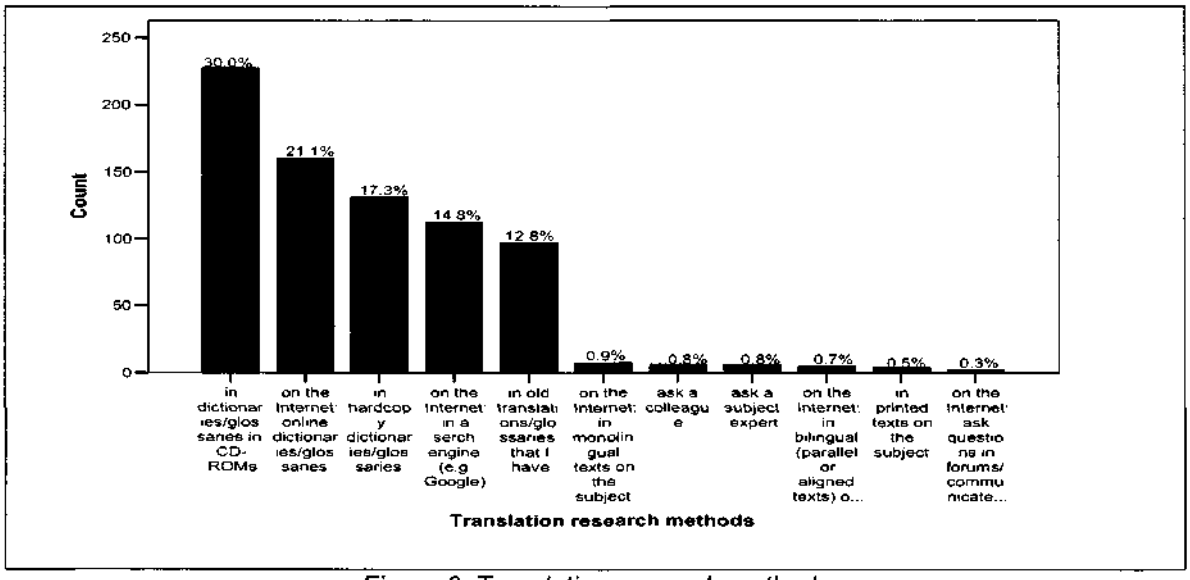

Figure 6: Translation research methods

4.3 TM usage rate and issues concerning low and no use

The percentage of individuals using a TM system was 82.5%, against 17.5% who did not use any TM systems at all. Figures 7 and 8 illustrate a detailed breakdown of user categories who use and those who do not use TM systems. In terms of working status, there is not a striking difference among different user categories, although it appears that company owners are slightly more likely to use TM systems, followed by company employees and then freelancers. This comes as no surprise, as both from findings of previous research (such as the 2004 LISA Translation Memory Survey) and from the fact that company owners have been the earliest adopters of TM technology, it can be said that they

are more open to TM use than any other group, being convinced about the cost savings and productivity gains deriving from the use of TM systems.

Of all different professions, translators constitute the substantial majority of TM users (39%). They are followed by project managers, reviewers, terminologists and a small representation of subtitlers, whereas interpreters and other translation professionals (such as graphic designers and DTP specialists) are more likely not to use a TM system, either because their job does not benefit from the use of such a tool, or because there are other specialised tools suitable for their particular job (see Figure 8).

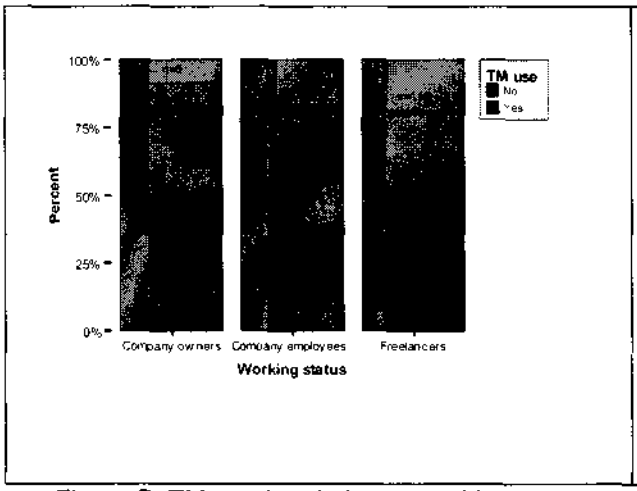

Figure 7: TM use in relation to working status

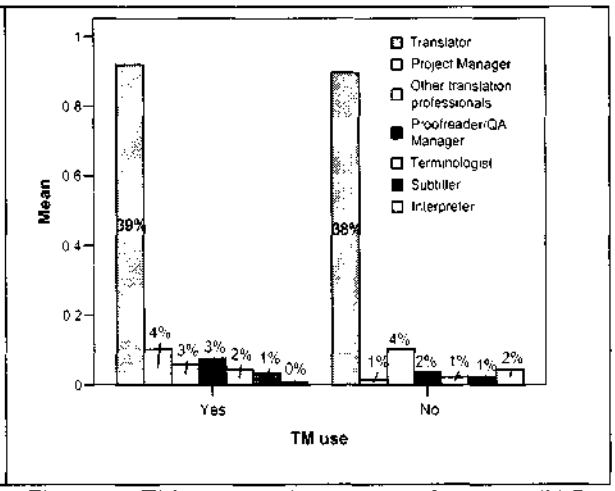

Figure 8: TM use in relation to professions (N.B. Interpreters account for 0.3% of the users)

The high TM usage rate (82.5%) revealed in this survey can be explained by two important characteristics of the sample unit. Firstly, the majority of respondents specialise in technical texts, and as already established in previous studies on TM use, there is a strong correlation between the particular text type and the use of TM systems. This is because technical content normally contains a great amount of terminology, standard expressions, simple sentence structure and a high degree of internal repetition; therefore the possibilities for content re-use are many. As Figure 9 also shows, those who specialise in technical texts are more likely to use TM tools, followed by those who specialise in financial and marketing content. Those who reported legal specialisation are also likely to use TM tools, but less than the previous groups, perhaps because although legal texts also contain a large amount of terminology, the sentence structure in legal texts is more complex and the level of internal repetition is normally lower. What is also evident from this figure is that, unsurprisingly, the TM usage rate drops for those who translate general texts (or have no specialisation) and for those who translate literature.

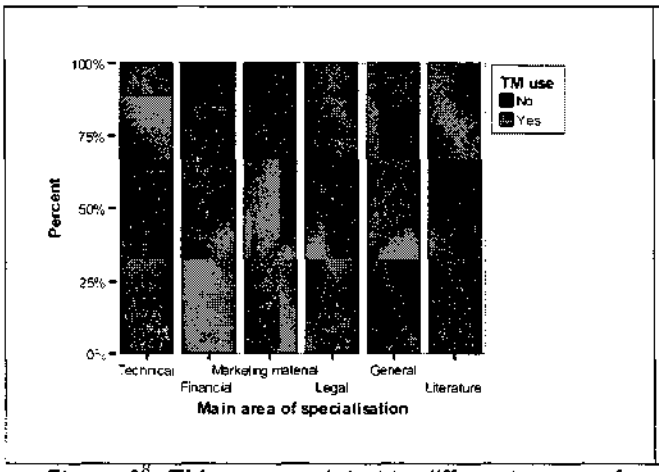

Figure 9[8]: TM use correlated to different areas of specialisation

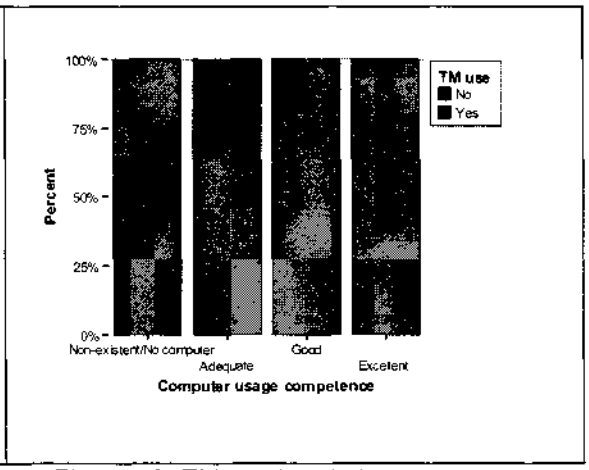

Figure 10: TM use in relation to computer usage competence

Secondly, another sample unit characteristic which has stimulated the high usage rate is the high level of computer usage competence that respondents have, which normally encourages the adoption of TM technology. In fact, as Figure 10 demonstrates, the more skilled the translation professionals are in the use of computers, the more likely they are to use a TM system.

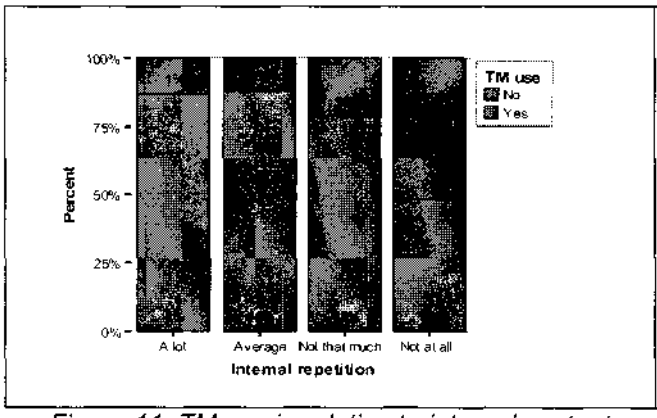

Figure 11: TM use in relation to internal content repetition

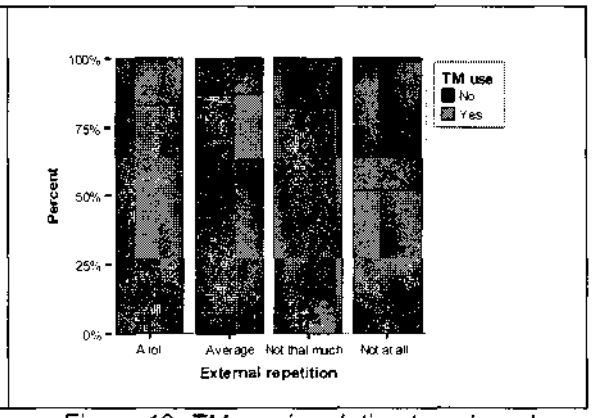

Figure 12: TM use in relation to external content repetition

An additional factor which influences the use of TM tools is the degree of internal as well as external repetition in the content. As Figures 11 and 12 illustrate, there is a relationship between high levels of repetition and high TM use, with the TM usage rate declining as the repetition levels decrease. Surprisingly, even the category of professionals who report not dealing with any content repetition at all are equally divided into TM users and non-users, although one would expect that TM tools would not be used as much in this category. This is probably explained by the fact that TM systems are not so dependent any more on content repetition in order to produce matches, but they are able to provide solutions either by deploying resources like termbases or integrated lexicons or by using effective matching techniques that can produce a match even for a segment that the system has not seen before.

[8] The percentages on the bars show the percentile distribution of respondents across the different categories. The sum of percentages should be 100%, which is the total number of respondents.

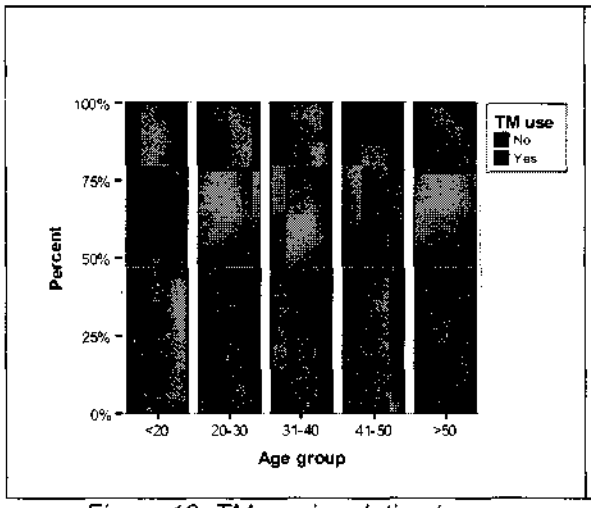

Figure 13: TM use in relation to age

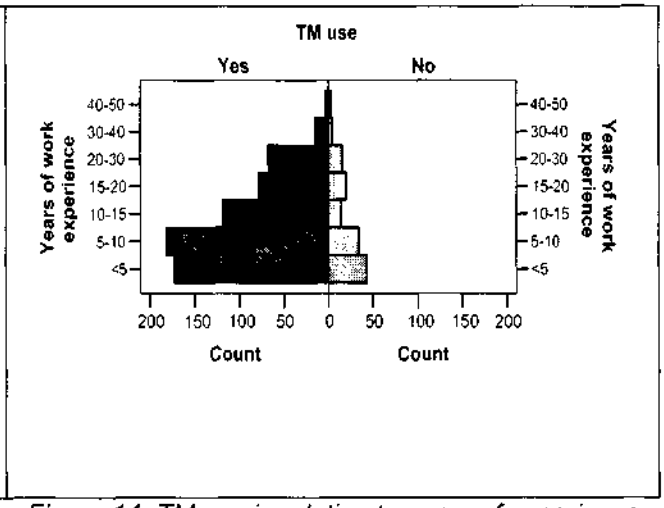

Figure 14: TM use in relation to years of experience

In an attempt to explore whether there is a relationship between TM use and age, the two variables were correlated (see Figure 13). The results did not bear any significant evidence of such a relationship, as no major differences were observed among the different age groups. The same applied in the correlation of TM use and the years of work experience, where again no relationship was found. As Figure 14 shows, there is a normal distribution of values for both TM users and non-users, which signifies that individuals with fewer years of experience use TM systems as much as those with many years of experience.

From the group of non-users, 89% have heard of TM tools, whereas 11% reported ignorance on the existence of such tools (see Table 2). For the non-users of TM systems, the commonest reason (28%) for not using such a system is because they believe a TM system is not a suitable tool for their particular type of work, whereas, surprisingly enough, 16% of respondents reported owning a TM tool but they have not been able to learn how to use it yet. Other common reasons reported were the high cost of purchase, the misconception that TM tools in general cost too much (this could indicate ignorance of the existence of free or low-cost TM tools), the dissatisfaction with the TM tools that the users tried, a lack of time and energy to search for a suitable TM tool and learn how to use it, and finally the TM tool's failure to bring any real benefit to the user's work after using one.

Reasons	Percentage
Not suitable for my work	28%
I have a TM software but I have not learned how to use it yet	16%
The one I want costs too much	14%
I've never heard of such tools	11%
I've been told they cost too much	11%
I have tried and evaluated one/several but didn't like any	5%
Lack of time/energy to find a suitable one and learn how to use it	4%
I was using one in the past but it didn't bring any real benefit to my work	3%

Table 2: Reasons for not using a TM tool

The percentage of non-TM-users who would be willing to try out or buy a TM system in the near future is 71%, which raises the hope of greater dissemination of the TM technology in the future. What we notice from the user profile of those who are planning to try a TM system (see Figure 15) is that all

company owners who are currently not using a TM system are planning to do so, as are the majority of company employees.

When these individuals were asked if they have heard of any TM system in particular and which one, Trados was mentioned by the majority (76%), followed by DejaVu, Wordfast, SDL Trados, SDLX, STAR Transit, MultiTrans, Passolo[9] and Omega-T (see Table 3). Several other TM tools were mentioned by fewer than 10% of respondents, such as (in order of popularity): CatsCradle, MetaTexis, Wordfisher, across, Alchemy Catalyst, LogiTrans, ForeignDesk, iLocalize, Logoport, AppleTrans, ENLASO Localization tools, Heartsome Translation Suite, Fusion Translate CMT, Lingobit Localizer, LogoVista, MemoQ, Multilizer, Visual Localize, WebBudget, TrAID, Translator's Intuition, Transolution, various internal tools, Cafetran, ProMemoria, SIMILIS, Transware Ambassador, and IBM Translation Manager.

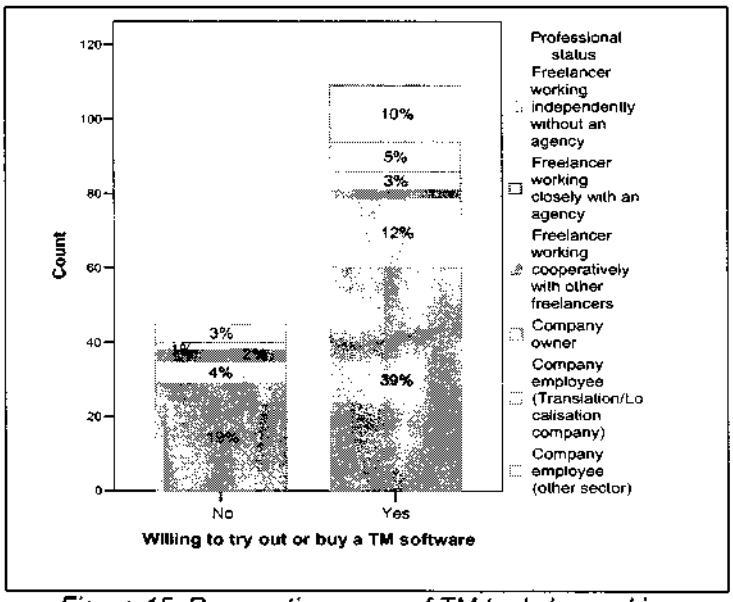

Figure 15: Prospective users of TM tools by working status

TM tool	Percentage[10]
Trados	76%
DejaVu	61%
Wordfast	51%
SDL Trados 2006	49%
SDLX	36%
STAR Transit	25%
MultiTrans	18%
Passolo	11 %
Omega-T	10%

Table 3: Top 10 most popular TM tools with non-TM-users

4.4 Perceptions & practices around TM usage

The survey also explored what prompted the use of TM systems for those who reported using one. The substantial majority of respondents reported using TM systems out of a personal choice (73%), as opposed to those who use the systems because they were imposed by the company they work for (or partner with) (20%) and those who use them because it was their client's request (10%). From Figure 16 it appears, not surprisingly, that company owners decide the use of TM systems on a voluntary basis, whereas the majority of company employees are forced to start using a TM either by their company or their clients. For the majority of freelancers, on the other hand, the decision seems to lie more in the assessment of personal circumstances, rather than in external requirements.

[9] Even though Passolo is a software localisation tool, it was mentioned as a known TM tool. Other software localisation tools are included in the analysis too, as they were mentioned in the respondents' preferences and ratings.

[10] Figures total more than 100% because respondents were able to select multiple tools.

In terms of the training provided on the use of TM systems, more than half of the users (51%) reported not having received any training (see Figure 17). In terms of those who have received training, 18% attended a short course or seminar, 12% received training from their company, 7% were trained on TM systems during an academic course and 7% received training from the TM software developer. Other sources of training were: colleagues, local software dealers and online user fora.

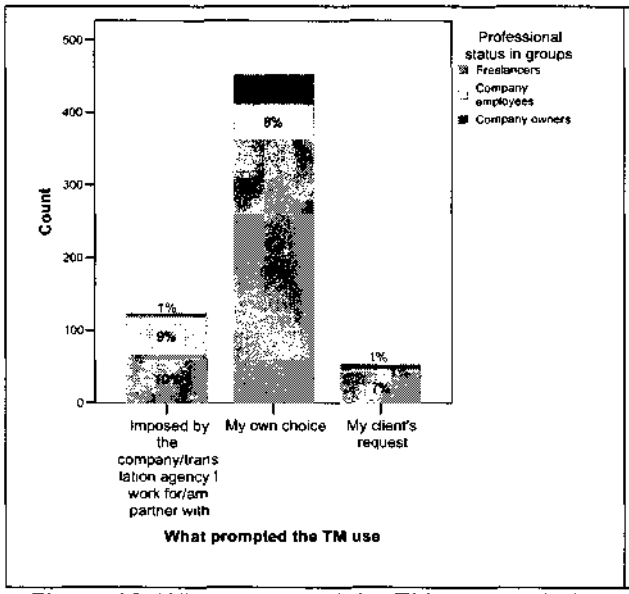

Figure 16: What prompted the TM use in relation to working status

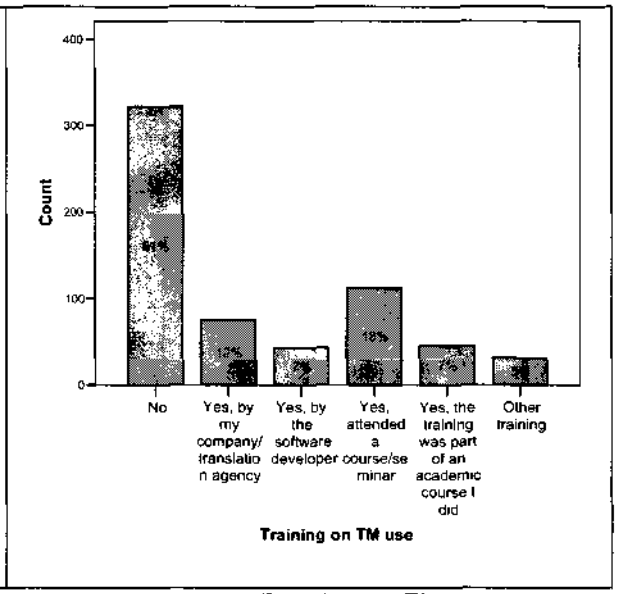

Figure 17: Training on TM use

The survey also examined the willingness of translation professionals to play a part in the development of TM systems. When they were asked if they would be willing to be involved in a TM system's development process (mainly through providing feedback on a TM product) 69% gave an affirmative answer, which appears to contradict some TM developers' argument that users in general do not seem keen on being engaged in the development of TM tools. Perhaps, their argument is sustained by failed attempts at engaging users, but these could be the result of a poorly thought-out method of engagement and not necessarily of the reluctance of the users to provide input. However, a number of users would not be willing to do so, mainly because, according to their responses, they either lack time for such an involvement, or they do not feel confident enough that an interaction with the developer could actually work and result in system improvements that would be to the advantage of the user.

The average length of TM usage is reported at 2-5 years, with a surprising 6% of users using TM tools for more than 10 years (see Figure 18). The level of usage reveals more positive findings for TM systems. A big proportion of the respondents (38%) reported using TM tools for 75-99% of their total content for translation, whereas 27% reported using the tools for all their content (see Figure 19). This result can be attributed perhaps to the fact that the majority of respondents specialise in technical texts, with an average or high degree of repetition, and they mostly work with standard file types (such as Word, Excel, plain text), all of which represent an ideal TM usage scenario.

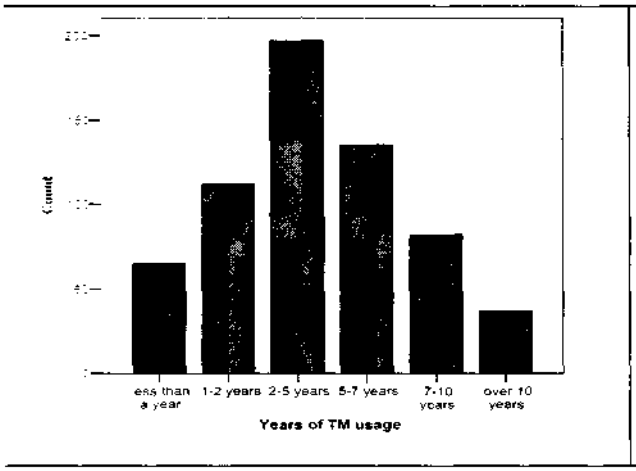

Figure 18: Length of TM usage

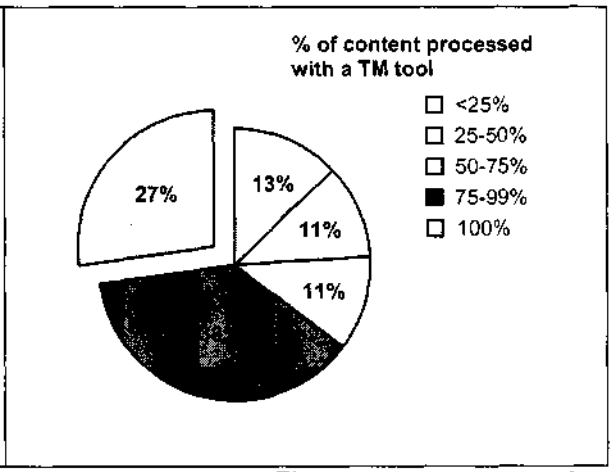

Figure 19: Level of TM usage (percentage of content processed with a TM tool)

TM users mentioned a number of reasons for not using TM systems for all their content (see Table 4). The most cited reason was the hardcopy documents which are still being used, as opposed to electronic documents. In fact, these types of files rank fourth in the file formats frequency (see Figure 5, section 4.2). Other important reasons were because some frequently used file formats (such as PDF and image files) are not supported by TM tools and because sometimes the text is too short to justify the setup of a TM process. One-off projects with low repetition, as well as texts on general subjects and literature also deter the use of a TM tool. A smaller percentage of respondents listed more reasons such as the complex layout of documents (with tables and/or embedded illustrations), the lack of training and inexperience, and the use of other tools such as DTP tools, terminology management systems and concordance tools. Perhaps some of these reasons (like the inability to support complex file formats and to deal with complex layout) could indicate the areas that TM developers and researchers need to focus when improving their TM systems.

Reason	Percentage
hardcopy documents	38%
the file format is not supported	28%
too complicated for short texts	18%
low repetition rate	18%
not suitable for my text types	15%
complex layout	5%
lack of training/inexperience	3%
use of other tools	2%

Table 4: Reasons for not using a TM tool for all content

In search of relationships between the length of TM usage and the percentage of content processed with a TM tool, some evidence is found to this end (see Figure 20). Up to the 10 years category, the level of TM usage increases steadily in proportion to the length of usage. This seems natural, since the more years of experience one has in using a TM tool, the more one understands its potential and feels more comfortable in using it for more content. The explanation may also derive from the fact that the translation resources residing in one's repository generally increase by the length of TM usage, offering a higher re-use rate for more content. Strangely enough, after 10 years of TM usage, we notice an insignificant decrease in the 100% usage level and a slight increase in the <25% level.

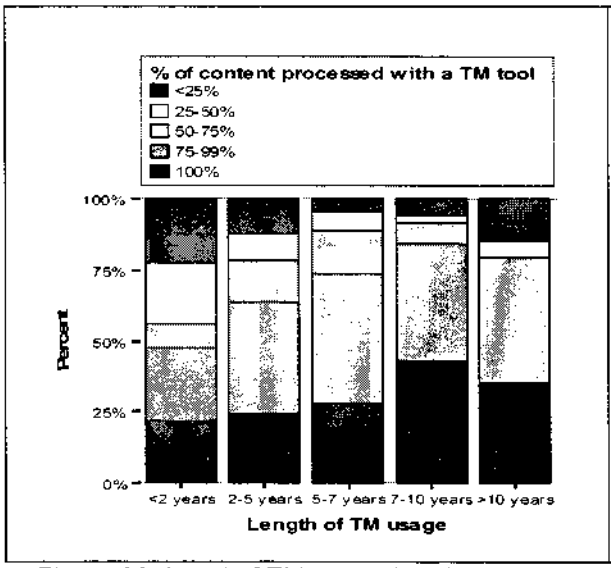

Figure 20: Level of TM usage in relation to the length of usage

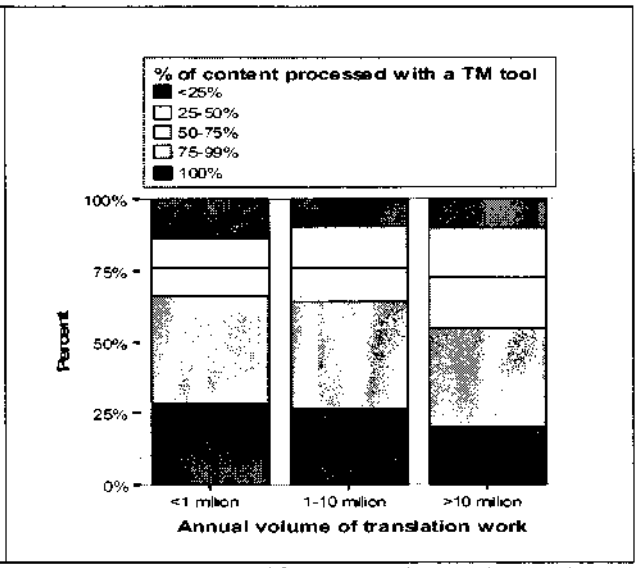

Figure 21: Level of TM usage in relation to the annual volume of work

In terms of the relationship between the volume of translation work and the level of TM usage, the high volume is not a good predictor of high level of usage. As it appears in Figure 21, no volume category presents a significant difference in how it affects the usage levels.

When users were asked why they voluntarily used a TM tool, the greatest proportion of respondents reported using a TM tool because it saved them time (86%), because the consistency in terminology improved (83%) and because the tool helped in improving the quality of the translation output (70%). Other perceived benefits were cost savings (34%) and the belief that TM tools offered the best way to exchange resources (such as glossaries and TM databases) (31%).

Regarding the most common tasks performed with the help of a TM tool, translation leads the way as the primary task, as it was mentioned by 96% of TM users. Terminology management (51%) and quality assurance checks (with regard to terminology -49%, consistency -49%- and completeness -43%) came next, followed by project management-related tasks (analysis for quotation and/or invoicing -43%), terminology extraction (24%) and checks for the client's proprietary needs (19%). Other less common tasks that were cited are: comparing and merging projects (16%), splitting TM databases to create project memories (14%), importing suppliers' TMs for in-house use (13%) and exchanging translation resources with colleagues (8%).

In terms of the more frequent ways to fill in a TM database, the significant majority of users (74%) stated that the database, initially empty, gets filled up as they translate. 51% reported aligning legacy translations with their originals in order to produce a database, 44% import their client's TM and 28% import their company's TM. Downloading pairs of parallel texts from the Web and then aligning them is an option for 15% of the users, aligning a ready-made bilingual parallel corpus is chosen by 12% and 6% use an external TM database that exists on a remote server.

4.5 Evaluation of TM systems based on users' preferences and experience

The majority (53%) of TM users use multiple tools, in comparison to 47% who feel a single tool is enough to perform their work. In general, TM users use from 1 to 8 tools, with an average of 2.21 tools. After comparing the average number of tools used with the length of time using TM, a striking trend is revealed. The number of tools increases as the user becomes more experienced in the use of TM tools over time (see Table 5). This is perhaps due to the fact that when users become more comfortable with the use of one tool, it is easier for them to learn and use more tools which may be better in some tasks :nan their first tool. Even more likely, the scope and type of translation work might change over time for some, creating new special needs which a single tool is unable to meet. Therefore, users are compelled to choose multiple tools, each one being fit for a specific purpose. In any case, this finding shows that users achieve a certain level of sophistication in the use of TM tools over time, which renders them more aware of the limitations and weaknesses of their current tools and drives them to investigate the TM market for new solutions to adopt, suitable to their needs.

Length of time using TM	Average number of TM tools
<1 year	1.46
1 -2 years	1.66
2-5 years	2.19
5-7 years	2.46
7-10 years	2.99
>10 years	2.76

Table 5: Average number of tools used in relation to length of TM use

User category	% using more than one tool	Average number of TM tools
Freelancers	52%	3.23
Company owners	54%	3.46
Company employees	58%	3.46

Table 6: Average number of tools used by user category

When the average number of tools used is examined with reference to user categories (as determined by their working status), it appears that company employees (58%) are more likely to use multiple tools, with an average of 3.46 tools, followed by company owners (54%), with the same average, and freelancers (52%) who use 3.23 tools on average (see Table 6). Company employees usually have access to a greater variety of TM tools compared to freelancers, mainly due to their cost, which explains why they use multiple tools more than the other user categories. Company owners follow company employees probably because the sample of this category mainly consists of small to medium-sized companies with a relatively small annual volume of work, which may not justify the purchase and use of multiple tools, again because of the cost involved.

The TM tools that are most widely used by TM users are listed in Table 7. Trados leads the field by accounting for 51% of the users, followed by Wordfast (29%), SDL Trados 2006 (24%), DejaVu (23%), SDLX (19%) and STAR Transit (14%). It is interesting to note that 3% of users have stated that they have built their own TM tool, or Word macros for translation, which shows again that a small but significant number of translation professionals can surpass the average level of usage competence and move toward developing their own solutions to their translation needs.

TM tool	Percentage of users
Trados	51%
Wordfast	29%
SDL Trados 2006	24%
DéjàVu	23%
SDLX	19%
STAR Transit	14%
Alchemy Catalyst	8%
Omega-T	7%
Logoport	6%
Passolo	5%
CatsCradle	4%
ENLASO Localisation tools	4%
Internal tool	3%
I have built my own tool (or Word macros for translation)	3%
across	2%
MultiTrans	2%
Wordfisher	2%
Heartsome Translation Suite	2%
Multilizer	2%

Table 7: TM tools more widely used

Other tools cited by less than 2% of the users were MetaTexis, MS LocStudio, AppleTrans. IBM Translation Manager, MemoQ, MS Helium. LogiTrans, ForeignDesk, Visual Localize. Cafetran, WebBudget, Fusion Translate, iLocalize, TrAID, Translator's Intuition, Transolution, CBG Transtool, Open Language Tools, TStream Editor Studio, LocFactory Editor, Kbabel, TStream Editor Studio, RC WinTrans 7, Lingotek, Prompt, Trans Web Express, Omega-t+ and Dr. Eye.

Comparing some of the most widely used TM systems with the user categories as determined by the working status, we come across a few interesting findings (see Figure 22). For instance, Wordfast appears to be used more by freelancers and company owners, but its use has not spread as much among company employees. Alchemy Catalyst, on the other hand, seems to be preferred by company employees more than any of the other two groups, and, in particular, it does not seem to be very popular among freelancers. Trados is preferred by company employees too. SDLX seems to be the choice especially of company owners, whereas STAR Transit is used more by company employees and freelancers, and slightly less by company owners. Finally, DejaVu presents a uniform distribution of usage among all three categories of users.

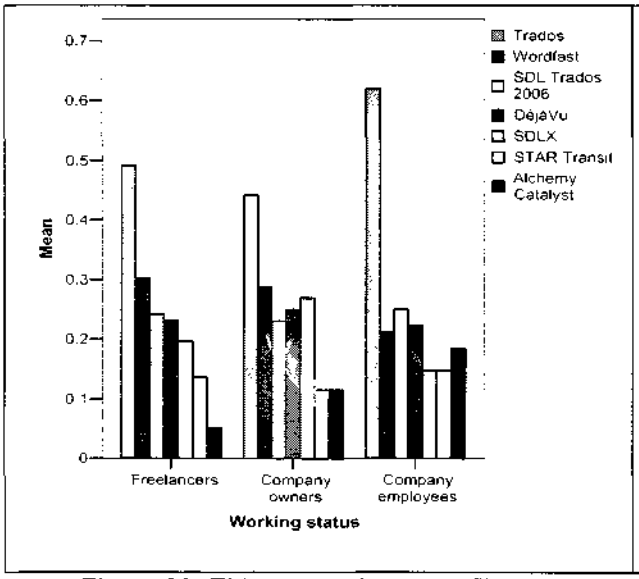

Figure 22: TM tools and user profiles (as determined by their working status)

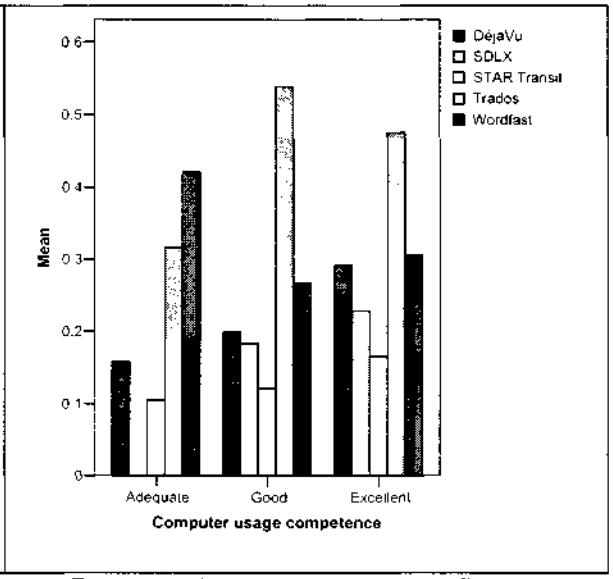

Figure 23: TM tools and user profiles (as determined by their computer usage competence)

Figure 23 examines a correlation of computer usage competence levels with TM tool preferences. The more striking observation is that SDLX seems not to have any users with adequate computer skills; instead the usage rate for this tool increases as the users' skills level goes higher. For Trados, we notice a higher concentration of users with medium-level ('good') computer skills, while DéjàVu and STAR Transit are both slightly preferred by high-tech users. Wordfast, on the other hand, represents an interesting case, by being significantly preferred by low-tech users.

Analysing further the TM tool preferences in relation to different user categories, there was an interest to reveal tool preferences for TM users who work mainly on platforms other than Windows. As Figure 24 illustrates, Wordfast is chosen by 27% of the users, followed by Omega-T, Trados, DejaVu, SDL Trados 2006 and the Heartsome Translation Suite. Although Trados, DejaVu and SDL Trados 2006 run exclusively on the Windows platform, we can assume that users employ those as a second tool on a second computer (either at home or at their workplace).

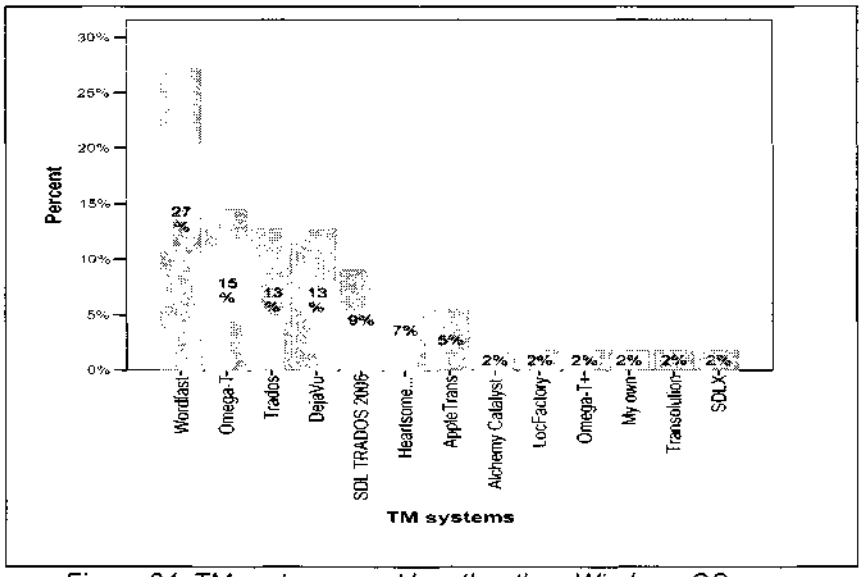

Figure 24: TM systems used by other than Windows OS users

TM users were also asked to indicate the single TM tool that they use more frequently, in order to address the cases where users had selected multiple TM tools in a previous question, and to provide a more valid evaluation of specific TM tools. Trados, as before, dominates in the users' preferences, accounting for 35% of the total TM users, followed by Wordfast, DejaVu, SDL Trados 2006, SDLX, Omega-T and STAR Transit (see Figure 25). It is worth noticing here that Omega-T, an free open-source TM tool with relatively few years on the market, has earned a sixth place in the top ten most frequently used TM tools. This could be attributed to the fact that the majority of users were freelancers, for whom the cost of a TM tool plays an important role in the choice of tool, but it could also mean that the TM tool has reached a certain level of quality that has allowed it to compete with other commercial tools and earn a high standing in terms of the users' preference.

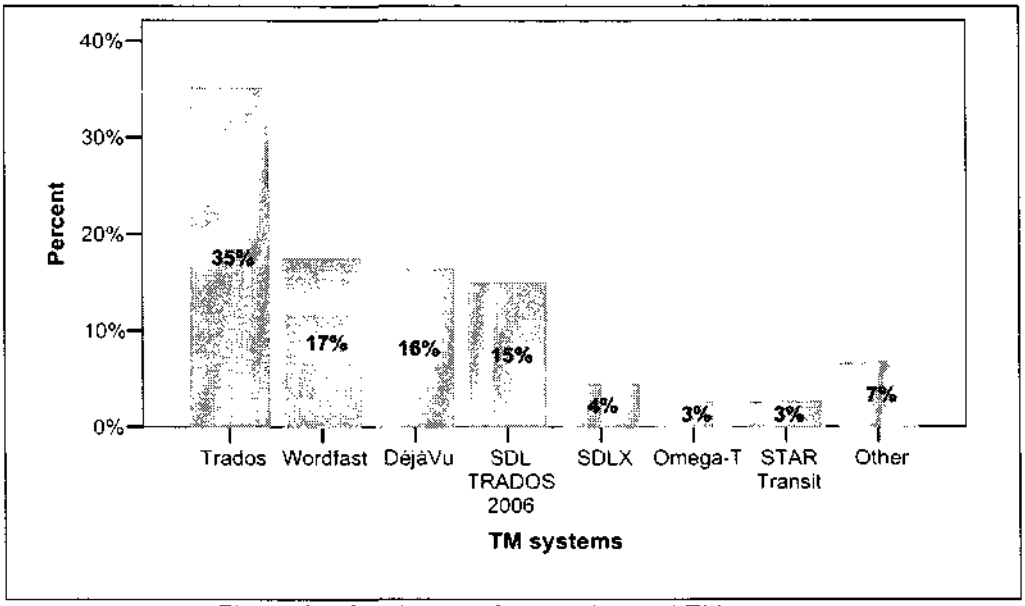

Figure 25: Single most frequently used TM system

Other TM tools cited, with a less than 3% concentration of users, were (in descending order of user concentration rate): MetaTexis, Heartsome Translation Suite, MemoQ, MultiTrans, AppleTrans, personal (self-developed) tool, across, Wordfisher, Alchemy Catalyst, Transolution, Omega-t+, LocFactory Editor, MS LocStudio, LogiTerm, Fusion Translate, the client's TM tool, LogiTrans, Logoport, Translator's Intuition and internal TM tools.

In order to find out which specific groups of professionals (in terms of their occupation) use each one of the four most frequently used TM tools, we broke down the total usage rate for each tool into the number of users in each user category (see Figures 26-29). From the general perspective, perhaps the most interesting findings from the general outlook of these figures, is that DejaVu seems to fall short in the representation of subtitlers in its user base, whereas Wordfast seems to be less preferred by project managers and other translation professionals (such as DTP specialists, graphic designers, web authors, etc.) among all user categories [11]. It might be possible that these preferences relate to the tools' lack of particular features needed by the specific categories of users.

[11] For interpreters, it is impossible to draw any valid conclusions, as the sample is very small (interpreters account for just 0.3% of TM users).

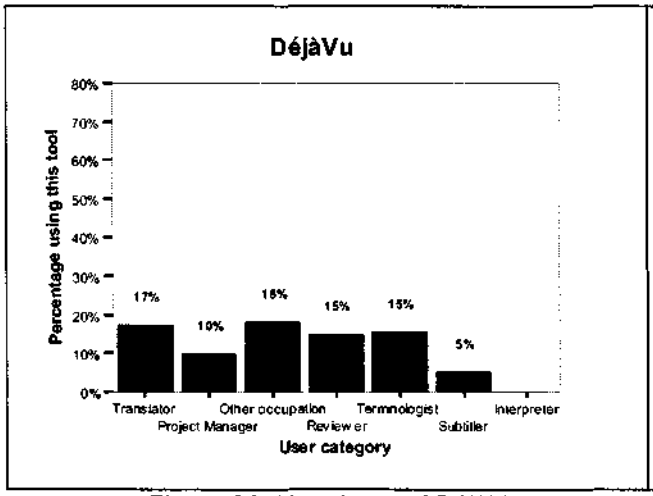

Figure 26: User base of DéjàVu

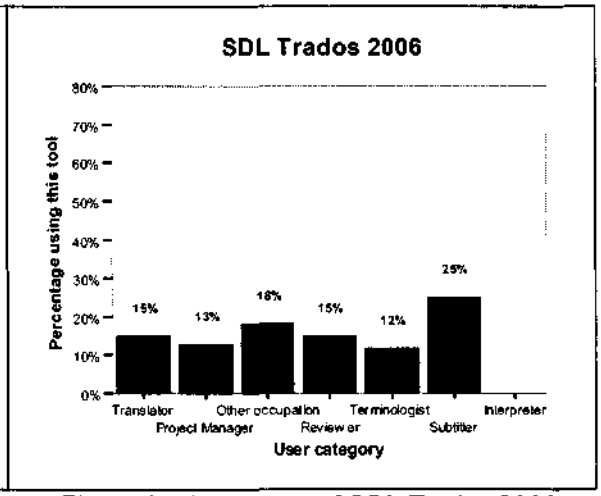

Figure 27: User base of SDL Trados 2006

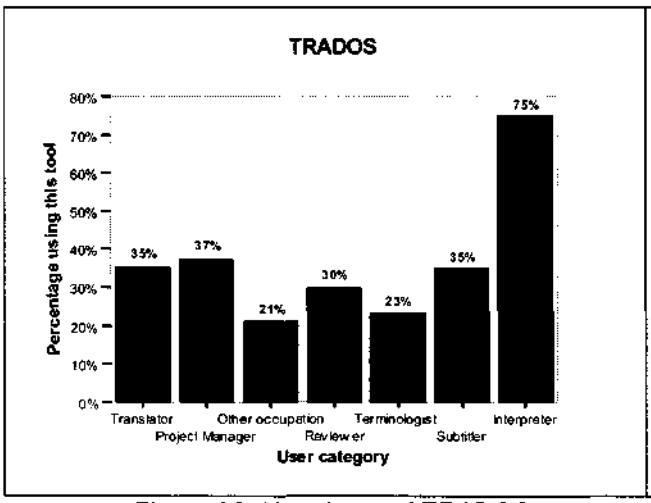

Figure 28: User base of TRADOS

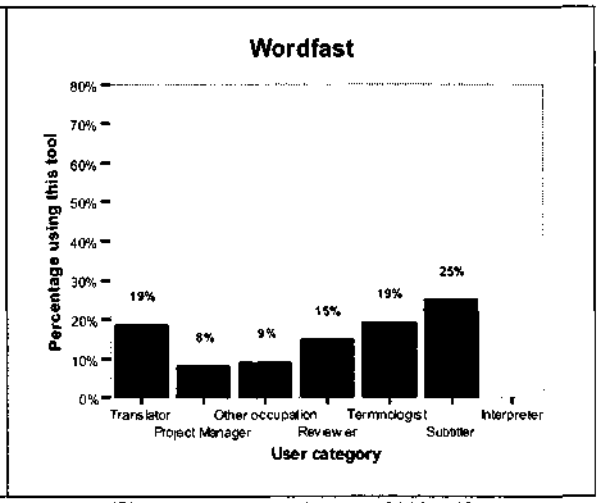

Figure 29: User base of Wordfast

TM users were also asked to rate the TM tool used most frequently, according to a number of functional and non-functional quality criteria defined below. The selection of criteria was based on the quality metrics proposed in the EAGLES framework (King, 1997) and on Höge's framework for evaluating translators' aids systems (Höge, 2002).

Functional criteria

- **Functionality:** as determined by the variety and relevance of features, which work as described in the product manuals
- **Efficiency in match recall:** as determined by the ability of the system to retrieve all available exact or fuzzy matches for a source segment
- **Efficiency in match precision:** as determined by the ability of the system to retrieve the correct exact or fuzzy matches for a source segment
- **Efficiency in speed:** as determined by the ability to search and retrieve matches quickly

Non-functional criteria

- **Reliability:** as determined by the frequency of errors and the occurrence of software bugs
- **Usability:** as determined by convenience and practicality in use through its user interface
- **Learnability:** as determined by the time and effort needed to learn how to use it
- **Value for money:** price to usefulness (as a tool) ratio
- **Customer support:** level of support received from the TM developer when needed

Table 8 shows the ratings attributed to Trados, DejaVu, Wordfast and SDL Trados 2006 by their users. Even though all TM systems used by more than 3% were rated, only the four most frequently used ones are presented here. The rating scale was: 1=Not satisfied at all, 2=Poor, 3=Quite satisfied, 4=Very satisfied, 5=Excellent!

Ratings in terms of:	**TRADOS**	**DejaVu**	**Wordfast**	**SDL TRADOS 2006**
Functionality	3.5	4.1	3.9	3.6
Efficiency in match recall	3.3	3.8	3.4	3.3
Efficiency in match precision	3.3	3.8	3.4	3.8
Efficiency in speed	3.7	4.1	3.9	3.7
Reliability	3.3	4.1	3.8	3.4
Usability	3.5	4.2	4.2	3.5
Learnability	3.3	4.0	4.1	3.4
Value for money	2.9	4.3	4.6	3.1
Customer support	2.5	3.7	3.7	2.6
Average rating	**3.3**	**4.0**	**3.9**	**3.4**

Table 8: Evaluation of four TM tools according to functional and non-functional criteria

From a general perspective, DejaVu seems to gain a higher rate of user satisfaction compared to the three competing TM packages. Wordfast comes second, followed by SDL Trados 2006, and Trados. It is worth noting that SDL Trados 2006 shows evidence of a marginal improvement against its previous version of Trados, in almost all areas, especially in terms of the efficiency in match precision. Wordfast, on the other hand, demonstrates higher satisfaction rate over its competitors in non-functional aspects like usability, learnability and value for money. Overall, the ratings suggest that TM users are generally satisfied with these tools, but there is obviously great room for improvement especially in the weaker aspects of every tool.

4.6 Future directions: visualising the ideal TM tool

The last section of the survey was addressed to all respondents, regardless of whether they used a TM system or not, with a view to gaining some insight into the way that translation professionals in general

(users and potential users) view some of the future developments regarding TM systems. As Figure 30 demonstrates, respondents have attached the greatest importance to the ability of the TM tool to handle text in embedded illustrations, to the compliance of the system with the TMX and SRX standards and to the ability to support PDF files. Spellcheckers for all languages are also considered important, as well as the TM tool's ability to display the context for the matches retrieved from the database (this feature along with a few more discussed here, is already provided by many TM systems, but it is worth showing how important it is considered by the majority of respondents). Other relatively important developments concern the ability of the system to construct and suggest fuzzy matches from all types of available resources in the database (when no match is found by the traditional methods) and the ability to view the full source text when translating a document, even when translating segment by segment. The preview of the full target text, the extraction of text from graphics and the support for all languages are also ranked as somewhat important features. Those which appear to be less important for most are the ability to search as you type, the support for project-specific directives, the ability to sort lists alphabetically and the integration of an OCR module. Abilities like plugging into a Content Management System and the possibility to run on multiple platforms are considered not very important developments. The former finding is explained perhaps by the fact that a small number of translation professionals work directly with Content Management Systems and are normally in-house employees (a user group with low representation in this survey), whereas the later finding derives from the fact that the significant majority of respondents are Windows users and all TM tools (with a few exceptions) run on Windows; so a platform-related development is not particularly of interest to them.

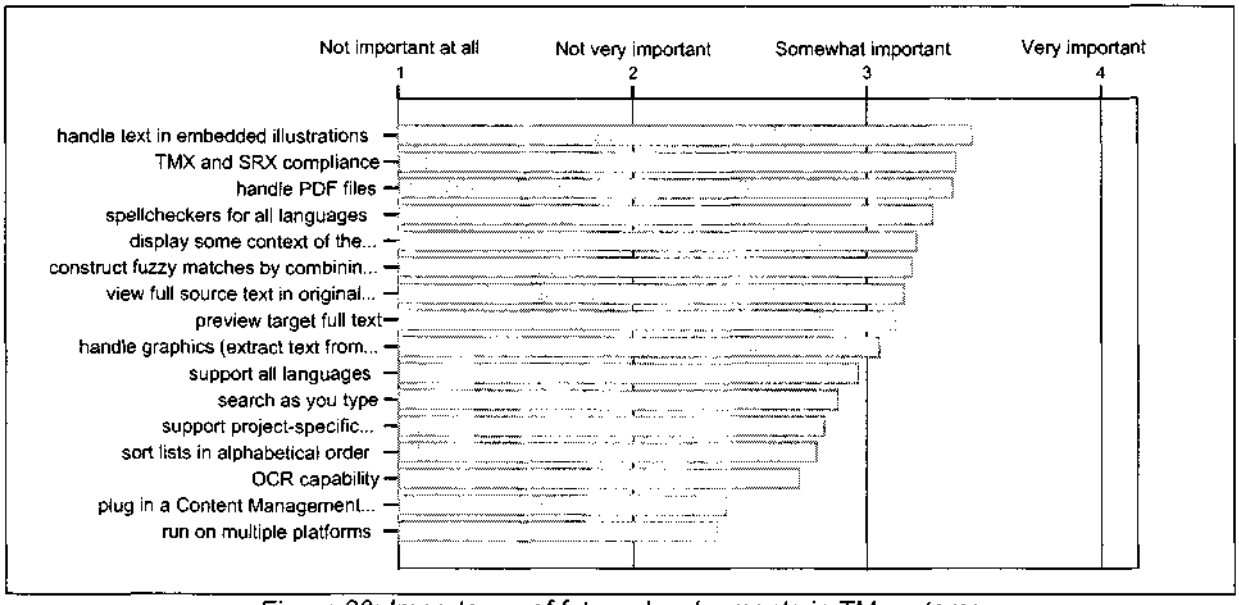

Figure 30: Importance of future developments in TM systems

The survey also explored users' (or potential users') preferences regarding the text processing environment that they find more practical to use (see Figure 31). 56% of respondents appear to prefer Microsoft Word, against 22 % who prefer using the TM application's text processing environment, 20% who indicated using both interchangeably and 3% who cited other various text processing software.

When asked whether they would prefer a single modularised TM application or multiple interactive applications (offered as a translation suite), each of which to be used for a specific group of tasks (e.g. terminology management, concordance, project management, etc.), 75% of respondents indicated preference for the first software solution, against 25% who chose the second. When this data was correlated with the computer usage competence levels (see Figure 32), there appeared to be no significant evidence that the two variables were related, as the majority of individuals with excellent computer usage skills still preferred a single application (like the majority from all other groups). This finding indicates the need expressed by the majority of translation professionals for simpler software solutions, instead of multiple applications which are more likely to add complexity to their work.

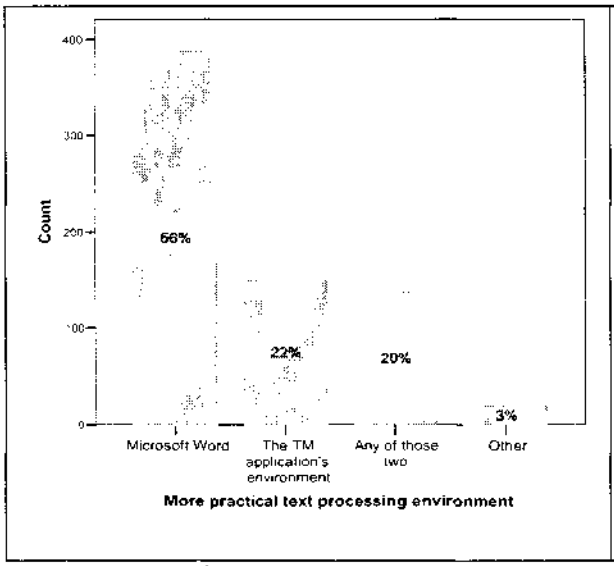

Figure 31: Text processing environment preferred

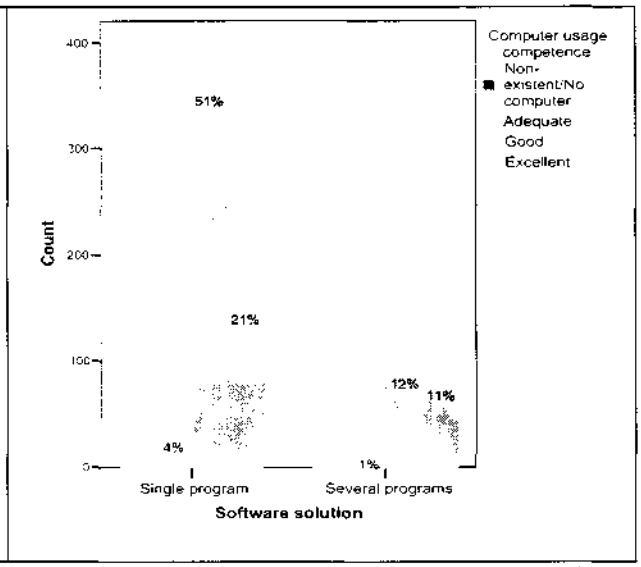

Figure 32: Software solution preferred in relation to computer usage competence

Given the growing interest among professionals regarding language resources (glossaries, dictionaries, corpora, translation memories, etc.) and their integration with TM systems, a few ideas were suggested through the survey and were put to the test (see Table 9). This aimed to identify new possible areas of research, justified by a substantial a level of demand.

Features related to language resources	**Very important** %	**Somewhat important** %	**Not very important** %	**Not at all important** %	**Don't know** %
ability to import dictionaries/glossaries from CD-ROMs	50.2%	30.2%	12.9%	4.6%	2.1%
ability to locate bilingual parallel texts on my subject on the Web (option available to specify authoritative sources only), download them, align them and fill in my TM database for future use or reference	43.5%	3.1.3%	16.3%	6.6%	2.3%
ability to locate webpages on the Web that contain glossaries for my subject and language pair, and bookmark them in a facility within the TM tool so that I can access them through my TM tool	43.8%	33.9%	15.0%	5.0%	2.3%

Table 9: Desirability of features involving language resources

As shown in the table, all three ideas suggested received a high importance rating from the majority of respondents. Translation professionals obviously supported the idea of unlocking the potential of language resources existing in CD-ROM, by bringing them into their TM database, thus being able to use a single and uniform repository of resources. They were also keen, as it appears, on the idea of some level of interaction between the Internet and the TM system. In effect, the ability to extract resources from the Web or to link dynamically to Web resources seems desirable for a significant proportion of the respondents.

5. Conclusions

The TM Survey 2006 examined the user's perspective in regards to the use of TM systems and produced a number of findings that allow us to draw certain conclusions about the state of the relationship between translation professionals and TM systems.

Overall, the survey found a high penetration rate (82.5%) of TM technology for its sample unit. The general characteristics of this unit, deriving from the background information on the respondents, are:

- the sample was made up exclusively of those with access to the Internet
- the majority were translators (90%) and freelancers (73%)
- most had a professional qualification relevant to their work (89%)
- the average length of work experience was 12 years, while the age is varied
- most rated their general computer usage competence as 'good' (at 64%), while 30% rated their competence as 'excellent'
- the greatest part specialised in technical content (61%) with high levels of content repetition
- all had access to the Internet (100%) and most had a fast (DSL) Internet connection (58%)

The above characteristics that synthesize the general profile of the survey's respondents bear a strong relationship with the high TM usage rate being reported in this survey, and confirm the findings of other studies on TM systems that have explored and demonstrated the factors which stimulate TM use.

Among the most important findings of the survey are the reasons for not using TM systems, and especially those reasons which reveal missed opportunities of use (such as the fact that 16% of non-users already have a TM system but they have been unable to learn how to use it). The promising finding concerning the group of non-users is their intention to try out or buy a TM system in the near future, reported by the majority (71%). The most popular TM tools for this group are Trados, DejaVu, Wordfast and SDL Trados, which could be something that testifies the successful marketing of these tools, since the respondents have not used them.

Another interesting finding is the willingness of the majority of respondents (69%) to provide feedback for a TM tool, thus getting involved in the development process of a tool that will meet their needs in a better way.

Regarding the length of TM usage, it turns out that users are moving towards longer periods of TM usage (with 6% of users using TM systems for over 10 years), and in terms of the percentage of the content processed with the help of a TM tool, a significant proportion of users (27%) use the systems for their entire content, whereas another 38% use the systems for 75-99% of their content. Both those findings are positive in terms of the implementation of TM technology and provide evidence of the users' realisation of the benefits deriving from the use of these systems.

The use of multiple TM systems by a large proportion of TM users (53%) is another interesting finding which invites a variety of explanations as discussed previously. The number of TM tools used appears to increase in proportion to the length of TM usage, and company employees seem to be the greatest user category which uses more than one TM tool (3.46 on average). The most widely used TM tools are Trados, Wordfast, SDL Trados 2006 and DéjàVu, each one appealing to different categories of users. Computer usage competence levels seem to bear a strong relationship with TM tool preferences; and so does the choice of the computer operating system.

The single most frequently used TM tool is Trados with 35% user concentration, followed by Wordfast (17%), DéjàVu (16%) and SDL Trados 2006 (15%). But although Trados appears to dominate the market, DéjàVu is the TM system which is rated highest in the users' evaluations. Users indicated as its main strengths the functionality, efficiency in speed, reliability, usability and price to usefulness ratio, which were all rated as exceptionally satisfactory. However, the evaluations overall asserted the need for improvements in all TM systems, and especially with regards to specific weakness exhibited by each system.

Finally, the survey provided a number of future directions in the development of TM systems and revealed preferences related to the design of a TM system. Respondents found the ability to handle text in embedded illustrations as the most important development, followed closely by the compliance of the system with the TMX and SRX standards (thus revealing an increased concern around resource portability issues), and the ability to handle PDF files (which are frequently encountered by translation professionals). The greatest proportion of respondents appeared to prefer working in MS Word, instead of the TM application's environment and the significant majority (75%) indicated a preference for a single TM application solution, instead of multiple applications offered as a TM tools package. Respondents also seemed to favour the idea of integrating Internet capabilities and Web resources into TM systems in order to increase their utility.

All the findings of this survey have helped to form a clearer picture of the relationship between translation professionals and TM systems. They are certainly open to further analysis, which may point

the way to new areas of research regarding the improvement of TM technology. We hope that this knowledge will enrich the pool of information available to TM researchers and developers, and will contribute to the advancement and higher deployment of TM technology.

Acknowledgements

I wish to thank my research supervisor at Imperial College, Mark Shuttleworth, for his generous support and feedback during the design of this survey and the analysis of results, as well as for his insightful corrections to this paper. I am also indebted to Joseph Eliahoo, of the IC Statistical Advisory Group, for his valuable advice on the statistical analysis of the data and the graphic representations. Finally, I am thankful to all those who contributed to the design of the survey, to those who promoted it, and, above all, to those who responded to it.

6. References

Gervais, Daniel (2002) "The Full-Text Multilingual Corpus: Breaking the Translation Memory Bottleneck" MultiCorpora R&D Inc. White Paper, April 2002.

Gow, Francie (2003) Metrics for Evaluating Translation Memory Software. MA thesis. University of Ottawa, Canada.

Grönroos, Mickel and Ari Becks (2005) "Bringing Intelligence to Translation Memory Technology". In proceedings of the International Conference Translating and the Computer 27, 24-25 November 2005, London (London: Aslib, 2005).

Griffin, Abbie, John R. Hauser (1993) "The Voice of the Customer". Marketing Science, vol. 12, no. 1, Winter 1993, pp. 1-27.

Höge, Monika (2002) Towards a Framework for the Evaluation of Translators' Aids Systems. PhD thesis. Helsinki University, Helsinki.

King, Margaret (1997) "Evaluation Design: The EAGLES framework" Konvens 97.

Kotonya, Gerald and Ian Sommerville (1998) Requirements Engineering. Chichester: John Wiley & Sons.

Levitt, Garry (2004) "New Developments in TM Sharing". Multilingual Computing & Technology. No. 67, vol. 15, issue 7.

Musale, Shailendra (2004) "Getting More From Translation Memories". Localisation Reader 2004-2005.

Planas, Emmanuel (2005) "SIMILIS: Second-generation translation memory software". In proceedings of the International Conference Translating and the Computer 27, 24-25 November 2005, London (London: Aslib, 2005).

Sommerville, Ian (2004) Software Engineering. 7th edition. Wokingham: Addison Wesley.

Zetzsche, Jost (2005) "The Translator's Tool Box. A Computer Primer for Translators." Oregon: International Writers' Group.