Translation memories and the translator

A report on a user survey

Dominik Schneider, Marcos Zampieri and
Josef van Genabith
Saarland University | University of Wolverhampton |
DFKI GmbH and Saarland Unversity

This paper presents a comprehensive study on the use of translation memory software by translators of different backgrounds. We designed a questionnaire that was completed by a pool of 723 respondents including professional translators, translation students, and lecturers in translation studies and translation practice. We analyse the results of the survey providing important information concerning user requirements, the most important features of TM software, users' perceived productivity, and market shares.

Keywords: translation memories, CAT tools, productivity, user survey

1.   Introduction

Computer-aided Translation (CAT) tools are an essential part of the modern translation workflow and the core of most widely used CAT tools are translation memories (TM). TMs allow translators to re-use previously translated segments improving the consistency of translations and increasing translators' productivity (Bowker 2005; Plitt and Maselott 2010; Zampieri and Vela 2014; Pal et al. 2016). TMs are particularly helpful in specialized domains in which repetition of technical terminology and textual patterns is rife, for example in legal texts, user manuals, technical documents, etc.

In spite of the importance of translation memory tools, to the best of our knowledge, only a small number of user surveys[1] on TM user satisfaction, requirements, and market shares have been carried out to date. This includes the work by

---

1. In this paper we use the terms "survey" and "user survey" interchangeably in order to describe how user perception is monitored by having participants filling out questionnaire.

Lagoudaki (2006), LeBlanc (2013), Zaretskaya et al. (2015), and Zaretskaya (2015) the latter focusing on the use of machine translation (MT) by professional translators rather than on TMs. In light of this, in this paper we present a comprehensive and up-to-date user study on TM usage and user requirements.

The main goal of this study is to investigate the extent to which today's translation memory tools meet the requirements of translators. One of the best ways to identify user requirements is through user surveys (Zaretskaya 2015) in which a set of questions is posed to a sample population of users (Maguire and Bevan 2002).

When evaluating software developed to increase productivity Zaretskaya points out that it is imperative to be aware of the users' needs as "user-centred design can increase productivity, enhance work quality, reduce support costs, and increase general user satisfaction". Particularly considering that auxiliary software is often not an enhancement, but simply required "in order to keep up with the industry requirements" and that "not all of the existing technologies were successfully adopted by professionals" (Zaretskaya 2015:1). A user survey such as the one presented in this paper can provide insights beyond mere sales figures and present the translators' reasons for purchasing and using a given tool and their overall satisfaction.

1.1   Related surveys

LeBlanc (2013) collected first-hand information by working together with professional translators at their workplace. The scope of his work, however, was limited both geographically and in terms of the diverse profile of the interviewed translators. Intentionally our survey is targeted at a very broad audience in order to cover as many backgrounds and to make it as representative as possible. In addition, it was not one of LeBlanc's goals to compare different providers and software products, but rather to analyze the use of computer-assisted translation tools in general. The survey by Lagoudaki (2006), on the other hand, represents the most comprehensive work in terms of participation and scope to date. However, given that it was published more than ten years ago, it does not cover the most recent technological developments in this field. Translation memory software has improved substantially in the last decade by incorporating linguistic knowledge to help more accurate segment retrieval (Gupta et al. 2016). Another important trend observed in the last few years is the availability of both translation memory and machine translation output in CAT tools (Alabau et al. 2014; Cettolo et al. 2014). As evidenced in recent MT shared tasks and competitions (Bojar et al. 2016) the results obtained by state-of-the-art machine translation software have improved with the emergence of neural machine translation (NMT) which allows MT output to be post-edited by translators with reduced effort. Examples of this

trend are MT plugins available for SDL Trados and web-based tools such as CATaLog (Pal et al. 2016) and MateCAT[2] (Federico et al. 2014).

Lagoudaki's survey is also no longer up-to-date in terms of market shares. After 2006, for example, *SDL* acquired its competitor *Trados* which had consequences both in terms of product development and market share. Similarly, translation memory technology has reached wide-spreadup-take, which might also have a significant impact on the overall situation. We added additional aspects to our survey which were not covered back in 2006, most significantly we asked participants which functions within TM software they consider most important which enables us to compare whether the current software satisfy the actual needs of the users. Since most translators use translation memories, the market for CAT software is expanding and software vendors compete offering new and better functionalities.

Finally, the work by Zaretskaya et al. (2015) and Zaretskaya (2015) is to the best of our knowledge the most up-to-date survey on the topic. It is comparable to our study in scope and methodology, however, as previously mentioned, it focused on the use of machine translation software rather than on translation memories. The aforementioned studies were carried out with different goals in mind and left a number of open questions related to user requirements, the functionality of TM software, and market shares that the present survey will investigate.

The remainder of this paper is organized as follows. Section 2 discusses issues in designing questionnaires focusing on best practices while taking the scope of our survey into account. Furthermore we present the profile of the 723 respondents who voluntarily participated in this survey. Section 3 provides the results of the survey with a detailed description of the outcomes. Section 4 concludes this paper highlighting the main finds of the survey and presenting avenues for future research.

2.   Methods

In this section we present the methods behind questionnaire design and the profile of the 723 respondents of this user study.

2.1   Issues in questionnaire design

To develop a consistent and informative user study we took the best practices in experiment and questionnaire design into account (Iarossi 2006; Krossnick and

---

2.   https://www.matecat.com/

Presser 2010). This concerns for example the type, structure, and wording of questions and answers as well as their order.

In this user survey we used both closed and open questions, which allow us to obtain both quantitative and qualitative information (Zaretskaya 2015). Creating 'good' questions is not a trivial task. Although there might not be a universal formula, Iarossi (2006) points out that "there is, however, a general agreement on what constitutes good and bad questions." Four criteria should be observed when wording questions: "it must be *brief, objective, simple* and *specific* (or BOSS)." This includes, for example, leaving out questions which are not relevant to the research as well as words not relevant to a question, avoiding complicated questions with more than three embedded clauses in one sentence (Peterson 2000), excessive question length (Payne 1951), and hidden questions (Iarossi 2006). In order to reduce negative impact on the leveraged data, the questions used in our user study were split into as many brief and clear questions as possible, using *skip* or *branch logic* wherever applicable and suitable.

Skip logic is an important feature in questionnaire design that changes the question a respondent will see based on how they answered the previous questions. We used introductory questions for each main topic of the user study and typically the participants are first asked whether or not they are familiar with a topic before receiving more detailed questions on that topic.

Iarossi (2006) points out that in multiple choice questions, the length of the list of possible answers also influences respondents. Options appearing at the beginning of a long list have a higher likelihood of being selected. This is called the *primacy effect* and to prevent it in our questionnaire, the range of options was either reduced or the question structure was changed. Another solution to cope with the primacy effect is the use of Likert-scale questions which ask respondents to provide an assessment (e.g. for a program function) on a pre-defined scale. Asking respondents to select the functions considered most important in a TM tool would have led to a long list of options, thus risking a 'leading' effect towards the options shown on top. Instead, this field was split into one Likert-scale question per function presented like 'on a scale from 1 to 5 how important is X?' For this type of question, there is no consensus on what is the optimal number of categories/scale points. Too few options provide insufficiently refined information, whereas too many make it difficult for the respondent to choose (Iarossi 2006). According to Iarossi (2006), experiments show that it is preferable to offer between five and nine categories. In our Likert-scale questions we opted to offer five categories.

As to the length of the questionnaire, Iarossi (2006) points out that "questionnaire length has a significant impact on data accuracy" and "longer questionnaires […] inevitably result in higher response errors" (Iarossi 2006:79). Therefore the

overall number of questions was kept to a minimum in order to keep the time required to complete the questionnaire below 5 minutes (on average).

2.2   Participants profile

The questionnaire was released in 2016 as a web-based form. It was send by e-mail to mailing lists in translation and posted in related forums and groups in social media. It was available for participants during a period of 3 months.

The questionnaire yielded a total number of 723 replies. In most questions, participants could choose whether they wish to reply or to skip to the next question, therefore the total number of respondents differs from question to question. Percentile shares and reply figures presented in the next sections are based on the number of replies given to each individual question.

The participants' profile of this user study was kept open and broad on purpose. We did so to receive input from people working on different kinds of translation-related professions. Calls including the link to the questionnaire were posted on several translation-related groups on social media, including the group of Bachelor and Master students in translation studies at Saarland University and the Universities of Mainz, as well as several international organisations. We distributed calls for participation in academic mailing lists which increased the number of respondents significantly, for example the German translators' and interpreters' association *BDÜ* promoted the study among its members, the *Translation Automation User Society (TAUS)* allowed us to post the questionnaire link on their *LinkedIn* page.

Due to the many different approaches described above, it was possible to reach translators not only with very different backgrounds, from students to translation agency owners, but also different nationalities and specialisations.

The actual distribution of respondents can be calculated from the first two questions which asked respondents' occupation and domains of specialisation. The final results for these participant profile-related questions are presented in Figure 1.

As can be seen in Figure 1, the largest group represented in the user study are freelance translators, either working (always or in part) through an agency or without any agency involved. In Figure 2 we present the domain of specialization of the respondents.
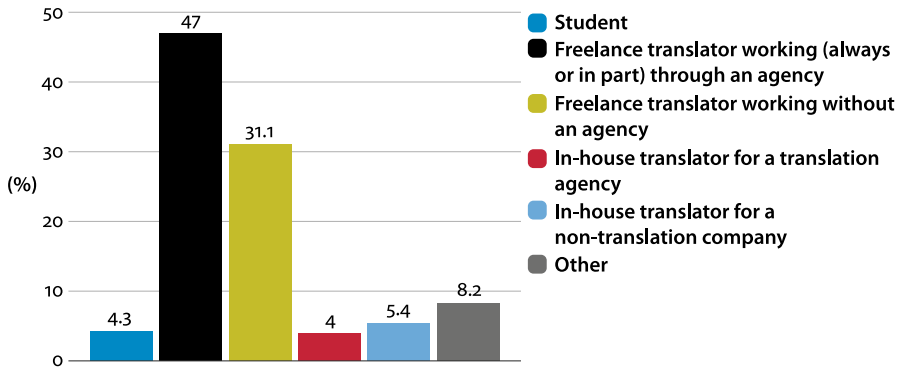
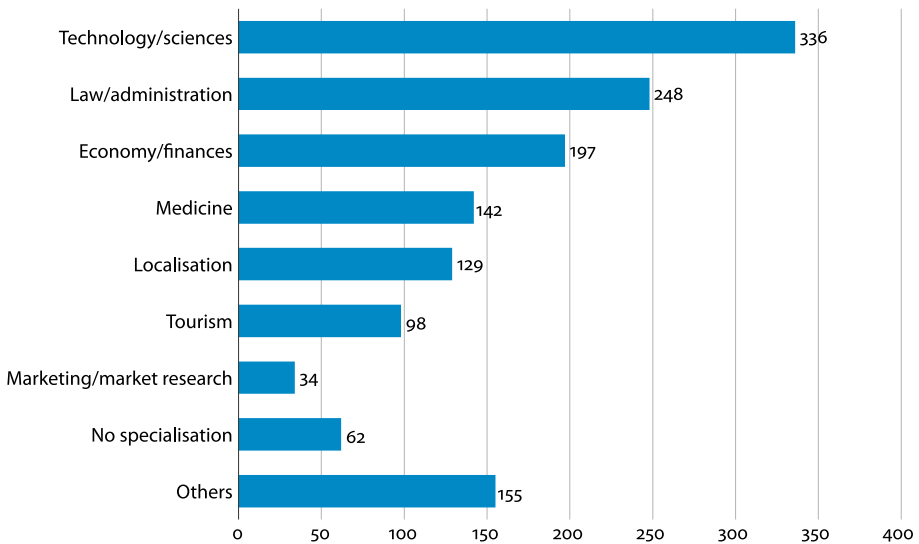**Figure 1.** Distribution of occupations

Figure 2. Distribution of specialisation domains (multiple choice possible)

Taking into account that many translators are specialised in more than one domain, the question asking for their domains has been designed as a multiple-choice question. In total, 444 of 723 participants (61.4%) have indicated multiple specialisations. Technology and sciences are the leading domains (336 translators indicated a specialisation in this field), but it has to be considered that they cover a broad spectrum of domains. Law/administration sectors follow (248 translators), economy and finances come in third place (197 translators). The diversity of possible specialisations and areas is emphasised by the large number of 'other' domains indicated (153 translators). These include: subtitling, games, academic writing, nursing, maritime/marine, research, (higher) education, cultural events, audiovisual, patents, and others.

3.   Results

In this section we present and analyze the results obtained in our survey. Section 4.1 presents results of software usage, Section 4.2 discusses market shares. Section 4.3 discusses how TM users perceive productivity gains and the importance of individual TM features for productivity and usability.

3.1   TM software usage

Following the first two occupation-related questions, participants were then asked whether they were familiar with translation memory technology or not. This question did not only serve the purpose of finding out how popular the TM tools are, but also to employ the aforementioned 'skip logic'. If the participant indicated that (s)he was not familiar with the technology or had not even heard of it, all subsequent questions detailing TM usage were skipped automatically. This way, non-qualified replies which would potentially harm the overall results could be prevented. Results are presented in Figure 3.
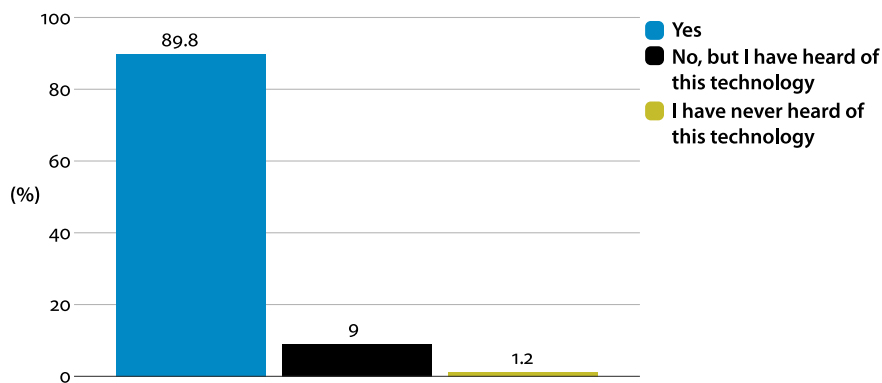

Figure 3. Familiarity with TM technology

The distribution yielded by this question unmistakably represents evidence for the widespread use of TM technology and CAT tools in general: an overwhelming 89.9% (649 respondents) are familiar with them, with 9% (65 respondents) at least knowing about their existence and only 1.2% (9 respondents) never having heard of them. This shows the importance that CAT tools have in the modern translation workflow. We decided to investigate familiarity and usage of TMs among students and we present these results in Figure 4.

These questions confirm that most academic institutions have successfully integrated TMs and CAT tools into their translation studies curriculum as all 33
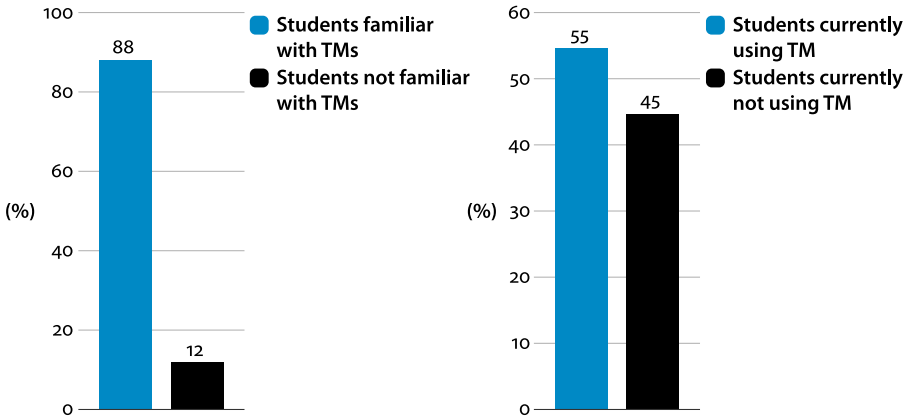
Figure 4. Comparison: Students' familiarity with and their usage of TMs

participating students with translation experience that took part in the user study had heard of the technology and 29 of them are completely familiar with how to use them. Usage distribution among students, however, is lowest amongst all kinds of occupation, with only around 55% of them currently using translation memory software for work. Considering their wide-spread familiarity with the technology, this may, at least in part, relate to the licence costs involved.

Although some of the tool providers offer free or low-cost versions of their software with limited functionality, acquisition costs might still be too high for student. Another conceivable explanation is that the students' translation volumes are too small to justify the software training periods and the time required to set up a database sufficient for increasing productivity when using a TM. Next we look at the familiarity of TMs among freelance translators. Results are presented in Figure 5.
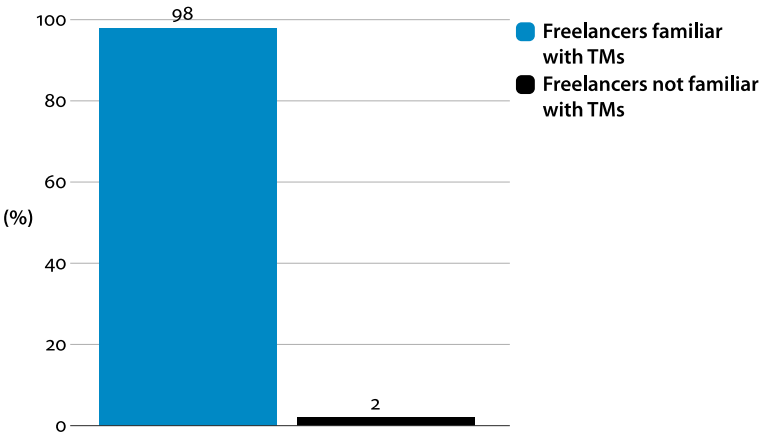

Figure 5. Agency-related freelancers' familiarity with TMs

The group of participants that displayed the most wide-spread familiarity with translation memories is the one comprised of freelance translators working with agencies (either exclusively or partly). Among the 340 translators who reported this occupation, 333 are familiar with TMs, which amounts to 97.9%. This might indicate that agencies promote the use of translation memories even more than the customers actually do.

The often mentioned correlation between translation memory technology and technical/scientific domains becomes evident in the familiarity with the programs. Overall, 48.4% of the respondents have indicated a specialisation in one such field, but of the 75 respondents who are not familiar with TM use, only 21 are technical/scientific translators, this amounts to just 28%. This seems to indicate that translators specialised in technological or scientific domains are much more likely to use translation memories than translators working in other domains.

Of the 649 respondents that are familiar with translation memory technology, 90.9% (590 translators) currently use TM software at their workplace. This means that only roughly 9% of translators have abandoned this technology once they had started using it. This might be caused by quite a number of reasons, such as a lack of IT skills or dissatisfaction with overall productivity levels when using the software. Licence costs may also play an important role, as the use of freeware software such as *OmegaT, WordFast Anywhere,* and *MateCat* is not yet widespread.

3.2   Market shares

Another important focus of our survey is how the translation market influence translators when choosing TM software. Do costumers require translators to use a particular tool? In Figure 6 we present the results obtained in our survey.

The question of whether customers require translators to use specific translation memory software led to a surprisingly even split, implying that there is no clear tendency on the market whether or not to use a specific provider/product. Around half of the respondents indicated that they are either often (40–70%) or in most cases (>70%) obliged to use a tool of the customer's choice, while the other half indicated that they have to do it only sometimes (10–40%), rarely (<10%) or even never. A more detailed view at the proportions, however, shows that the largest number of replies (199 or 30.9%) account for 'in most case", signifying that almost one third of translators are used to more than two out of three translation jobs received are tied to specific TM software.

Participants were subsequently asked which translation memory tool(s) they are currently using. As it was likely that the majority of translators works with more than one software, this question was designed as multiple-choice. The results are presented in Figure 7.
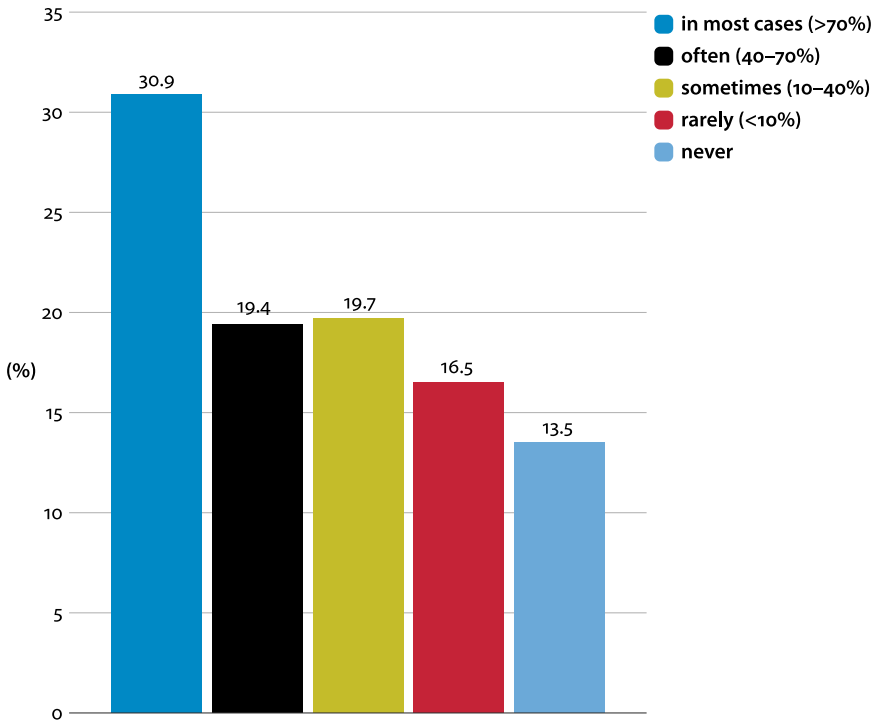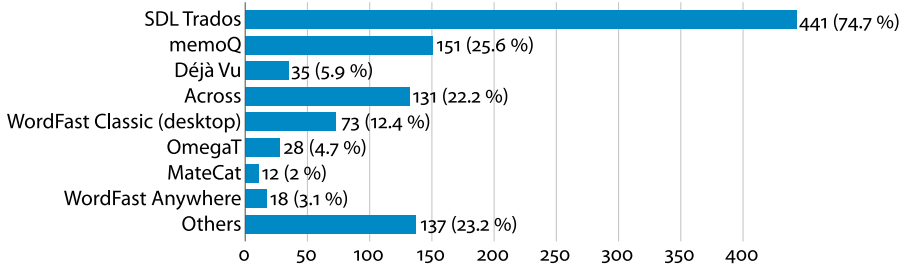
Figure 6. Enforced TM software usage

Figure 7. Current market shares of TM applications (multiple choice possible)[3]

In our survey almost three out of four translators are currently using *SDL Trados* which confirms this tool as the market leader. Although *Trados* stands at the top by a large margin, tools such as *memoQ* (25.6%), *Across* (22.2%), and *WordFast* (12.4%) are also being used by a significant number of translators.

---

3. In order to consider those translators working with mutliple TM programs alongside each other, this question was designed as multiple-choice. The result is therefore displayed as a bar chart.

Furthermore, 257 respondents or 58.3% *Trados* users indicated that they use more than one TM tool at their workplace. In Figure 8 we present the percentage of users that use one or more TM software.
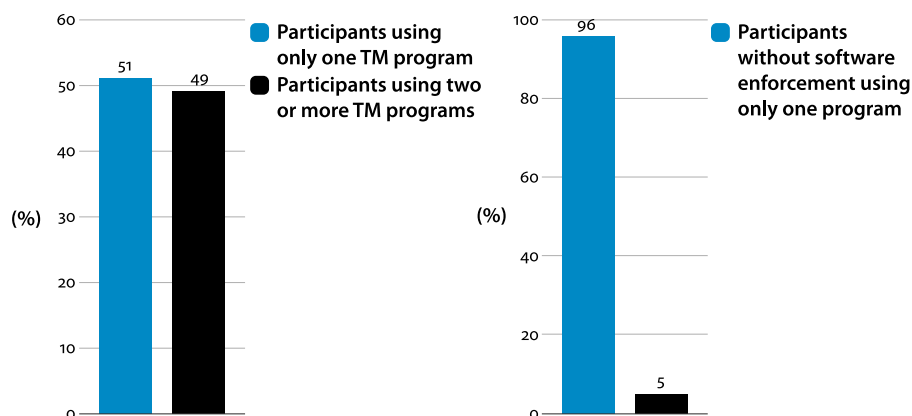
Figure 8. Comparison: Complementary use of multiple TM programs – overall and for those translators whose clients do not enforce the use of specific TM software

287 (48.6%) participants use two or more translation memory programs at their workplace. The indicated combinations of providers and products are quite varied, without any clear tendency discernible. Many translators even use three or more programs simultaneously. The obligation to use specific TM tools has a direct influence on how many of them the translators use. Among the 87 respondents who indicated that their clients never enforce specific software to be used, only four (4.5%) use more than one translation memory program. All others in that group either do not use any TM at all, or have settled on only one particular tool.

The fact that most translators use more than one TM software can be related to customer- or agency-side enforcement of certain software or, as pointed out by Zaretskaya (2015), it "might be a sign that there is no such software on the market that could satisfy everybody's needs, and many professionals have to resort to various tools to be able to take advantage of all features they like" (Zaretskaya 2015).

Among the translation students who took part in this user study, only one does not use *SDL Trados.* This is not surprising and confirms that academic institutions tend to use the market leader software to meet market demands. The large proportion of 'other' tools is also worth mention. These include both other desktop-based programs such as STAR *Transit* (NXT) and specialised programs, for example the software localisation tools SDL *Passolo* and Alchemy Software's

Catalyst as well as online-based alternatives, e. g. *Memsource* and *WordBee*. There-fore, the impressive share of 23.2% is much more due to the large number of *different* tools on the market which all bring about relatively small user bases rather than *one* 'other' tool representing serious competition. Online translation memory services such as *WordFast Anywhere* and *MateCat* (3.1% and 2% respec-tively), however, have so far had a fairly small impact in the maket, a situation that might change in the coming years.

Next we look at the influence of customer enforcement on the market leader, *SDL Trados*. In Figure 9 we present the proportion of *Trados* users among respon-dents that indicated customer-side software enforcement.
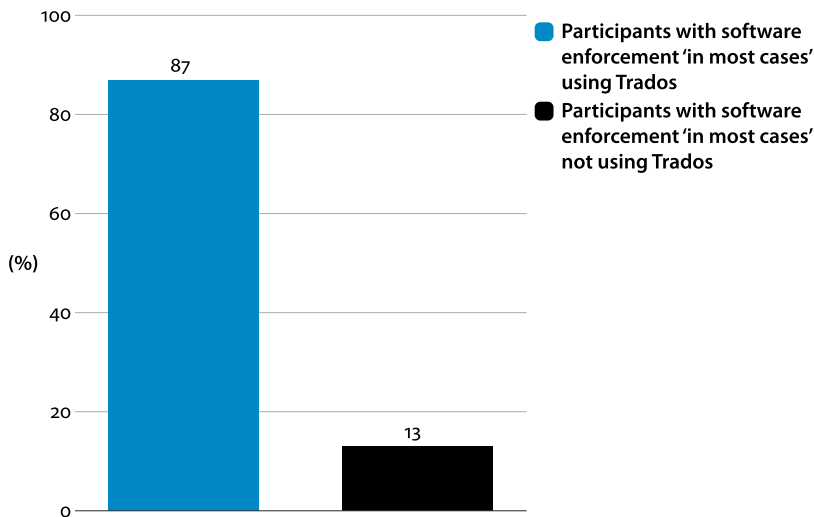

Figure 9. Trados usage among translators with software enforcement 'in most cases'

Among translators whose clients request the use of translation memory soft-ware 'in most cases', 171 of 196 respondents indicated the use of SDL *Trados*. This amounts to 87.3% and exceeds the general market distribution of the program by almost another 13%. This shows that when customers request the use of a partic-ular tool, it is most often *Trados*.

Next we look at whether the market distribution discussed previously is based on popularity rather than on the quality of tools. In Figure 10 we show the percentage of users which compared products before starting to use a given tool.

Only a slim majority (51.2% or 301 respondents) tested multiple systems before making a decision. As the internet makes product comparison possible for everyone and as all brands provide free evaluation licences, this can mean that many translators are led to use a particular translation memory tool through the
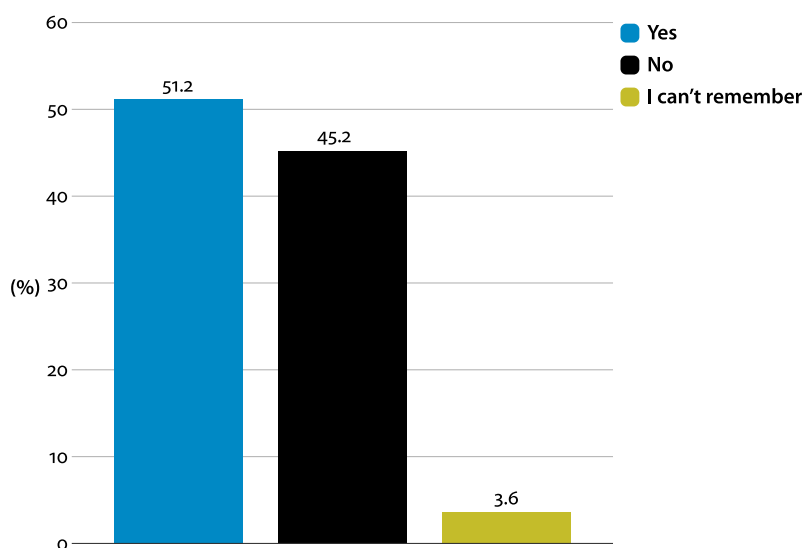
Figure 10. Awareness and examination of competing products before committing to the software

enforcement of their employer and/or their customer(s). Once again taking into account the high licence costs usually involved in the purchase of such software, it is highly improbable that users deliberately went without any research on available products and just purchased the first tool they found out about.

3.3   Perceived productivity gain

Next we look at how translators perceive possible productivity gains with the use of TMs. In translation process research, it is generally agreed that TMs increase productivity in almost all domains (Plitt and Maselott 2010), but to the best of our knowledge, translators' assessment on this matter has not been significantly explored. In Figure 11 we present the percentage of respondents that indicated performance gains when using TMs.

The result indicates that translators do perceive productivity gain by using TM software. A total of 87.9% (518 respondents) believe that they either work slightly more or even much more quickly and effectively with the help of the software. The majority (56% or 330 responses) actually account for 'much more quickly and effectively', thus emphasising the value of TM use. Only 13 participants (2.2%) believe that they work more slowly than before, with the rest either indicating that their performance has not changed noticeably (32 respondents or 5.4%) or that they do not know.
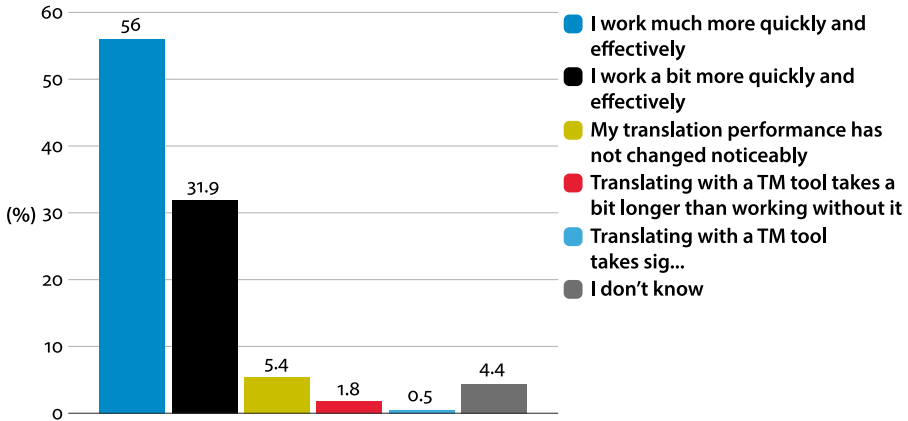
Figure 11. Perceived productivity development when using TM

3.3.1 *Evaluating individual TM software features*

In this section we look at individual features often included in translation memory software. User study participants were asked what degree of importance they attribute to each feature. The results allow us to investigate what are the most used and the most essential functionalities of TM software. For this purpose, participants were asked to rate the importance of TM functions on a scale from 1 (unimportant) to 5 (very important), with the middle point 3 being equivalent to 'neutral opinion'. Figure 12 presents answers on text alignment functions.
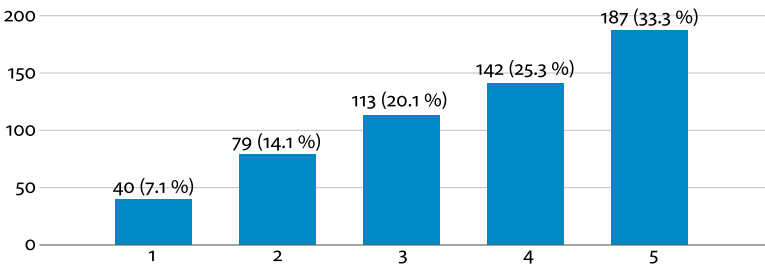

Figure 12. User-perceived alignment tool importance

As previous questions revealed a significantly large group of translators who either use multiple systems or have switched from one program to another in the past, alignment might indeed become an important part of any TM program suite. This latter hypothesis turned out to be true: the rating receiving one third – and thus most – votes was 'very important' (187 or 33.3%), with the diagram taking an echelon form towards 'unimportant', which was chosen by only 40 respondents (7.1%). However, 231 votes (41.3%) falling into the categories

'unimportant', 'relatively unimportant' and 'neutral' prove that indeed, there is a significant share of translators who do not require an alignment tool for their translation work.
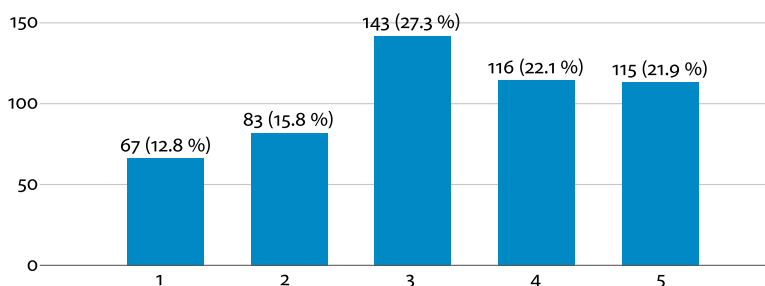

Figure 13. User-perceived bilingual term extraction importance

Automatic term extraction is already included as a function in most translation memory programs – however, this usually remains a monolingual procedure: in that case, source language terms are only exported and need to be complemented with a translation manually by the user as (s)he works his/her way through the document. Only few programs currently on the market offer a bilingual terminology extraction functionality and furthermore, this is often reserved to standalone programs designed specifically for terminology management (such as SDL *MultiTerm Extract*) instead of translation memory program suites. Nevertheless, participants were asked how much they care about this feature. It was to be expected that a certain quantity of respondents would consider it unimportant, since most probably, only a few of them possessed software already capable of this function. Surprisingly, this latter group only constitutes a minority: only a total of 28.6% (150 respondents) rated the feature either 'unimportant' or 'rather unimportant'. On the other hand, 44% (231 respondents) perceive it as either 'rather important'or even 'very important'. This shows that translators are indeed open for innovation and automation. Of course, it has to be noted that the largest, individual share of votes was received by the 'neutral' option (27.3% or 143 respondents). Nonetheless, the user study results to this question clearly mean that bilingual terminology extraction represents a worthy investment for program developers.

Project management has found its way into most TM-based programs and tools. As it seems to be a natural part of the translation workflow today, one could expect broad support and appreciation from the translation community. Once again, the user study's results proved otherwise: almost one third of the respondents (172 or 31.5%) consider project management to be 'unimportant' and make it the dominant opinion. Adding the votes for 'rather unimportant', one even arrives
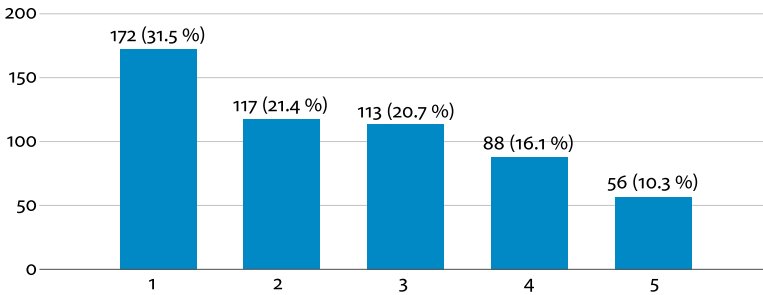
Figure 14. User-perceived project management functionality importance

at the majority of translators (52.9% or 289 respondents) who do not require the feature. Only a mere 10.3% (56 respondents) consider it 'very important'. Project management features are crucial to language service providers, but our survey shows that this is not true for translators using TM software.
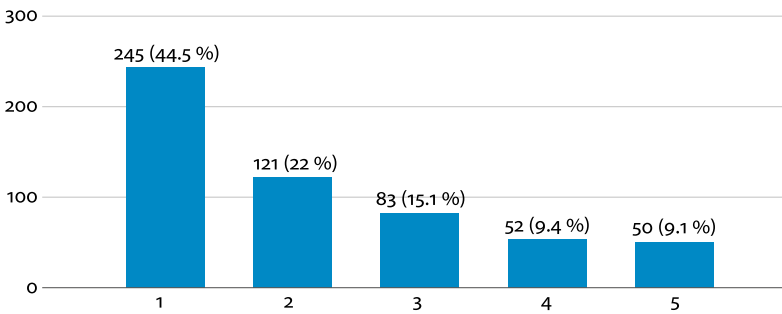

Figure 15. User-perceived machine translation interface importance

One of the most controversial subjects within the language service community is certainly machine translation, thus it was to be expected that the translators' opinion would not be unanimous, to say the least. The results show a clear tendency: almost half of the respondents (245 or 44.5%) voted an interface for machine translation output to be an 'unimportant' feature within TM programs. This is yet another setback for all those who were hoping that automatic machine translation would soon be widely accepted within the community – as the result shows unmistakably how strong scepticism towards it still remains. Besides the fact that many translators might fear for their very existence in the market being replaced gradually by computers and/or their salary further reduced the more TM and MT are used, our interpretation of these results is that the quality of the machine translation output remains dissatisfactory to translators and still requires much post-editing effort.

Only 9.1% (50 respondents) seem to have fully accepted MT as part of their workflow by calling it a 'very important' feature, with another 9.4% (52 respondents) at least considering it 'important' It is thus standing to reason to conclude that recent tendencies towards establishing machine translation might require strong progress in order to succeed, while it remains unclear whether translators' scepticism will ever be overcome. Students' replies to this question, on the other hand, might indicate a change to come in the next number of years: only one of the participating students considered an interface for machine translation output to be outright 'unimportant', so the perception of future generations might differ greatly.
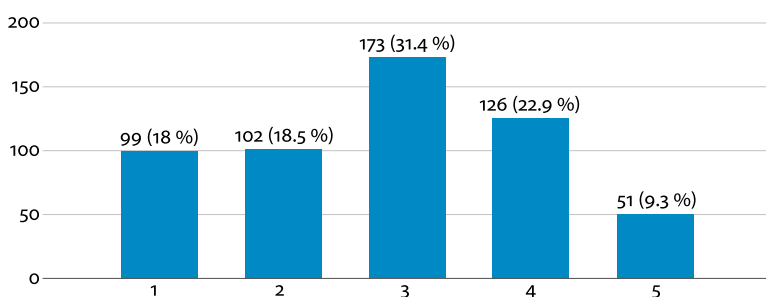
Figure 16. User-perceived contemporary design importance

Whenever human and machine interact, it is usually not only about mechanics but also importantly about presentation. The software, being the layer where both worlds come in contact with each other, needs to make the computer's resources available via an interface whose layout can have a significant impact upon the translator's performance. Many providers have taken efforts to produce an up-to-date and intuitive design – but do users appreciate this trend? One thing is certain: the result to this question was not really to be expected, as translators seem very focused on pragmatics alone. While 'neutral' was the most selected option (173 respondents or 31.4%), only 51 participants (9.3%) deem the contemporariness of the interface to be an essential factor of translation memory programs. This means that only roughly one third (32.2% or 177 respondents) care about this topic, while the rest is either indifferent or considering it negligible.

The user study results on machine translation output presented in Figure 15 might have given the impression that translators are very sceptic towards letting a computer do 'their' work. The question about quality assurance functions can therefore be seen as an insight into their opinion when the computer rather does 'rudimentary' tasks like checking for errors of all kinds. Here, distribution is very clear: the majority of the respondents (312 or 54.6%) have voted this
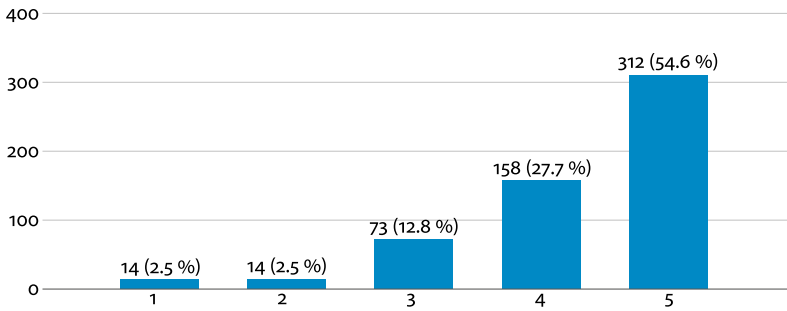
Figure 17. User-perceived quality assurance importance

field to be 'very important' when working with translation memories. Only 28 participants (5%) in total consider it either 'rather unimportant' or 'unimportant', with 12.8% (73 votes) being neutral. Therefore, more than four fifths of all respondents (470 or 82.3%) share the habit of using the computer's resources for this category of tasks. This might be a consequence of the fact that most word-processing programs – the most famous of the field being Microsoft's *Office* suite – have been using basic quality assurance tools such as spell and grammar checkers for many years. Just like most computer users, the strong majority of translators will be familiar with and accustomed to these applications, which makes it easy for TM program developers to introduce and maintain these as an established part of the CAT workflow as well.
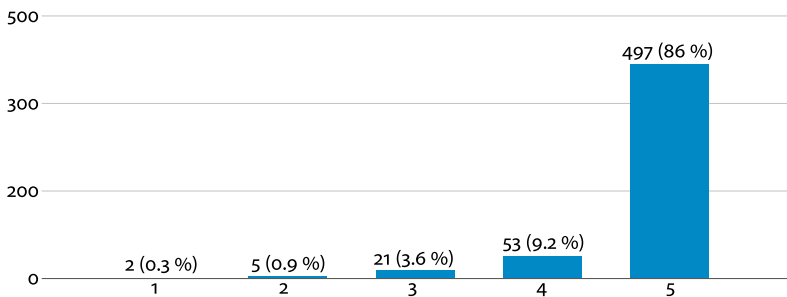

Figure 18. User-perceived concordance search importance

If one might be inclined to call the result of the previous question clear, then the participants' opinion on the concordance search function would exceed anyone's expectation even more. The outright majority of participants (497 or 86%) consider concordance search to be 'very important' for their translation workflow, with everyone else except a mere 1.2% (7 participants) being at least 'neutral' towards this functionality. Therefore, this question yields the most obvious and unanimous result of the entire questionnaire. If we compare this

to the users' opinion on more 'automated' applications such as machine translation, one might be inclined to conclude that translators appreciate a maximum of transparency and traceability as well as a certain support of 'manual work' in the tools they use. The 'thinking' should still be done by the humans themselves, with the machine doing rather the tedious processes in the background than handing everything to the user 'on a silver platter' without further interaction required.
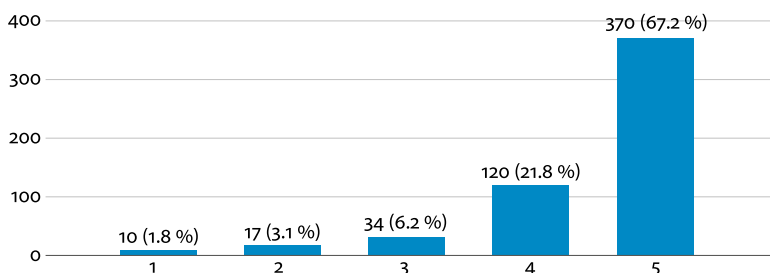

Figure 19. User-perceived simultaneous database use importance

When the question about the simultaneous use of multiple translation memory databases in one project was set up, we expected only a minority of participants to welcome this function. However, the organisation of resources appears to be an essential point in the translators' workflow. In order to bundle those resources related to particular domains or topics, it seems that most translators organise these in individual files: a two-third majority (370 respondents or 67.2%) considers the simultaneous usability of multiple TM databases a 'very important' feature and another 21.8% (120 respondents) at least 'important', while only 27 participants (4.9%) do not care about this possibility at all. Considering that some providers have not included this feature or, respectively, have only enabled basic options, this is a surprising development. For example, only very few programs allow the user to update several databases at once with new segments (usually, only one file can be accessed in 'write mode') which represents a definitive, time-consuming obstacle within a workflow that is based on productivity, effectiveness and speed.

With the background knowledge about transparency and traceability stemming from the question about concordance search discussed earlier, the requirement for context information to facilitate translation decisions seems a logical consequence. The distribution of replies to this question confirms this assumption: a total of 84.7% or 480 respondents appreciate this kind of information for the results presented by their translation memory software, with the slim majority (51.9% or 294 participants) even considering it 'very important'. Only 5.3% (30 respondents) don't consider this feature important. Once again, it comes as quite
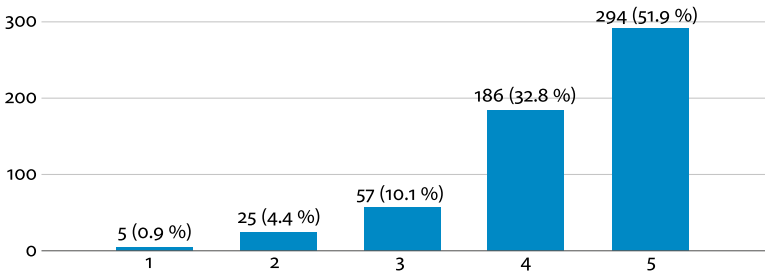
Figure 20. User-perceived context information importance

a surprise that only a few of the providers have truly elaborated on this functionality: only few programs managed to take things one step further by not only displaying the name of the file which includes the given translation memory result, but also offering to show the segment within the actual document, thus allowing to check surrounding sentences (the actual 'context'). Not surprisingly, translators consider background information important, as picking one suggested translation 'out of context' is a very risky procedure since the topic of the original document might be completely different and thus unsuitable. Trying to reach the same point of certitude and quality assessment by using other tools such as concordance search might be possible after all, but compared to the successful implementation of context information, that alternative would be much more time-consuming and tedious.
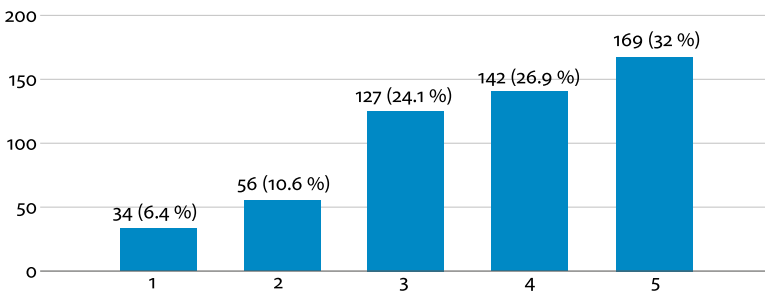

Figure 21. User-perceived report and log output importance

Digital reports and logs were typically features used by rather advanced users but the results obtained in our survey indicate that more users are finding these features important. For example, 169 respondents or 32% found these logs and reports 'very important'.

3.4   Qualitative analysis

From the 723 respondents, 157 took advantage of the opportunity to comment in free text on the translation memory program(s) they are using (or have used in the past). The topics found in their comments are manifold and address a multitude of different issues and functionalities; therefore it makes most sense to present just a summary of what was mentioned most frequently.

Criticism starts already relating to the compatibility with multiple operating systems: although Microsoft *Windows* clearly dominates the market, several participants have called for *Linux* and *Mac* versions, which, for example, the market leader SDL *Trados* does not offer at all. The high subscription/licence fees for most software have also been subject to much criticism from translators. The lack of proper inter-compatibility between different TM programs has also been pointed out by many participants.

One of the topics mentioned several times by participants is the integration of speech recognition/dictation in the functionality range of the TM programs, or at least a better compatibility with software for this purpose – users have particularly criticised SDL *Trados* for not addressing their requirements in that area. As most programs on the market offer a broad range of supported file formats, filter compatibility was not often mentioned as a point of criticism – but when this was the case, it was usually related to the PDF format either not being supported or not working well. Better implementation of XML has also been asked for several times. On a more advanced note, users have also been asking for OCR (optical character recognition) being implemented in TM software in order to use it with documents only available in printed-out format. Similarly, deficiencies in the alignment function have been mentioned many times, again particularly relating to *Trados*. However, according to the comments received, none of the programs currently available on the market seems to offer satisfactory performance in this category. The same consensus applies to tag management: it seems that no program has found the 'right' solution to take away the additional work required to handle and properly insert these.

Participants have reported bugs for almost all programs, for example *memoQ* (e.g. installation problems) or *Trados* (when changing segmentation, when using *MultiTerm* or the comment function). Many respondents demand that developers test their programs more thoroughly before a version is officially released to prevent the amount of bugs and errors increasing even further.

What the respondents have criticised most (by far), is that translation memory software has been enhanced by so many features that their complexity has risen to a level that makes training absolutely necessary, as developers seem not to have done much to maintain intuitiveness in their programs, at least for

basic features. The changes and new features included with new program versions seem to make it impossible for many translators to memorise all the options and settings required, which leaves many wishing for a simpler and more efficient system. In this context, it should also be mentioned that the multitude of menus and buttons required to cover all functions included with the software has made the use of a mouse basically obligatory, while many users ask for a rather keyboard-based approach which would allow time saving through shortcuts. The lack of flexibility in that respect for most programs has been criticised many times, just like the fact that the pre-set shortcuts differ from program to program. Customisation is also an often mentioned term relating to the general interface of the applications: many developers have implemented a certain degree of possible personalisation, but many settings still remain fixed. Furthermore, the wish for scaled-down versions has been expressed, particularly in terms of a reduction of the 'project management' emphasis which currently takes the focus away from the actual 'translation' work and forces the user to click his way through wizards and dialogue boxes until translation work can eventually be started. In these processes, many programs create a multitude of different files in different places of the computer's storage, causing additional maintenance work to keep the system in a 'healthy' state. Frequently, respondents claim in their comments that they prefer earlier program versions due to those being simpler, easier to learn and reduced to essential features.

Rather 'basic' tasks not working properly have been mentioned several times in the users' criticism, such as the joining and splitting of segments, which might be necessary for example when they have been badly aligned. It seems that particularly *Trados* does not implement this function perfectly yet.

Context information for results suggested by the software is obviously scarce in many cases – which is what many users mentioned as well. It is not always clear for them where exactly the suggested term/translation came from, so it proves difficult to judge how reliable the suggestions are.

Relating to terminology, users demand both an improvement in reliability and correctness rate of the function, as well as a proper implementation of bilingual term extraction, which was, for example, not included with any of the programmes reviewed earlier.

Although most systems boast compatibility with allegedly all languages of the world, many translators reported problems encountered with some languages that obviously have not been implemented completely. For example it was mentioned that Chinese was not displayed properly in *Across*, which made the respondent switch to *MetaTexis* where it worked fine. Similarly, problems with Thai seem to occur across many programs. A rather particular issue that, still, has been reported by several respondents is the strict assignment of language variations to

a translation memory database. The example coming up most frequently for this case are the UK and US variations of the English language. As all programs clearly differentiate between these two, a TM database for English (US) cannot be used for a document in English (UK) language, even though it would yield many high-quality matches.

Concordance search – which is considered a highly important function by most participants as presented earlier – also does bring with it some imperfections according to multiple comments. It seems that particularly *Trados* is criticised for its non-ergonomic concordance interface and for costing the user lots of time when switching between source document and target document search. Furthermore, many users demand additional filters to further specify which (part of the) translation memories are searched through, as otherwise large databases increase the response time of the search. According to one respondent, *Trados*' concordance search function also has had problems with languages that use many long compound words such as German or Finnish for many years – yet this still seems to remain unresolved.

Although many translators have expressed their appreciation for quality assurance (QA) features within translation memory software, apparently most programs do not deliver a satisfactory performance in that respect. Whether it be the fact that some of those QA checks strongly depend on the language pair used (for example, a check for consistent capitalisation cannot be used in German as nouns are always capitalised, as opposed to English for example), that a numbers check yields many false positives where a number in the source text has been spelled out in the target text or that the spell checker simply does not 'know' all words used in the translation, quality control still seems largely inferior to what many translators know from other word-processing software. Since agencies, employers and customers still seem to require these QA checks to be performed and since some TM programs force translators to click their way through all issues found in the document before the target language document can be exported, quality assurance proves to be a task consuming a significant amount of time in many cases.

Despite the scepticism of many translators towards internet-based functions (see the following paragraph on users' opinions towards online TM tools), several participants have called for an even more 'cloud-based' approach of the programs on the market. This might make it a very difficult task for program developers to satisfy the demands of all translators, as opinions on this topic seem to diverge strongly. In the same vein, some participants wish for a more 'universal' search function which also includes popular online resources. Similarly, some machine translation supporters wish for further development such as the possibility to let the software suggest edits to the machine translation output

in order to make it fit the actual source segment better. Others do not refuse MT, but would prefer a clearer differentiation (or 'hierarchy') between the results found in the actual translation memory from those that have been output by a machine translation engine.

As the user study has shown, many translators work with multiple translation memory programs simultaneously, which also allows them to compare software directly. Therefore, many users desire some programs to implement functions they found to be present elsewhere, such as the live target text preview, which has been implemented in the main interface of *memoQ* for example, but is only an 'optional' feature in *Trados*.

The possibility to comment on programs does not mean that only criticism has been uttered: many participants have also seized the opportunity to express their satisfaction with changes and developments made in recent years. For example, it has been acclaimed that some of the tools now have a function included which allows the import of the customer's changes to the created translation (revision) back into the program, such as *memoQ*'s 'monolingual review'.
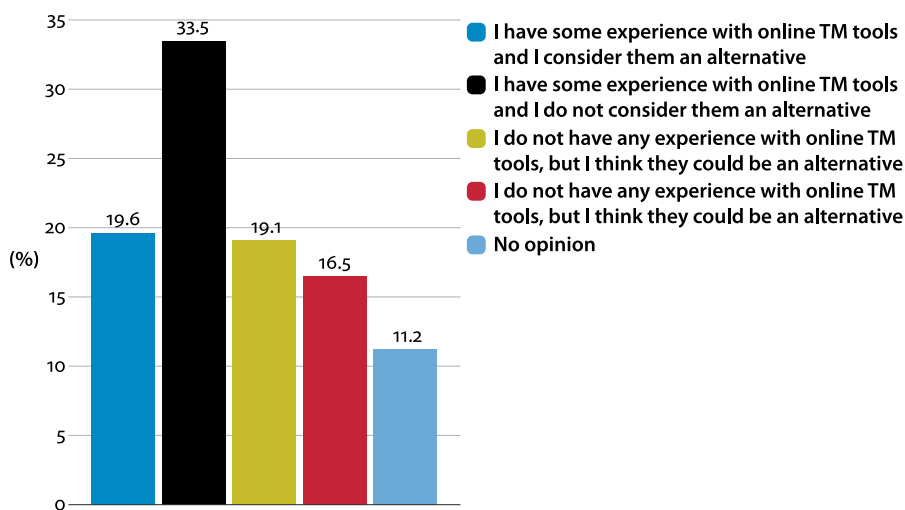
3.5   Online-based TM tools

Very interesting responses and comments on online-based TM applications could not only be found from within the questionnaire, but also from the remarks we have received from several translators who contacted us directly after completing the questionnaire. Many of them wanted to further explain their choices in the survey, particularly relating to privacy concerns when using online applications. They were sceptic about how their uploaded, private resources and databases would be used by the programs, let alone them being published and made accessible to others without their consent. Most translators consider their own translation memory databases and terminology lists – possibly gathered and built over decades – an important part of their competitive assets and fear losing competitiveness if other translators in the market could also take possession of those resources. This is just one of possibly many dimensions that exist to the question whether or not online TM tools such as *WordFast Anywhere* or *MateCat* are considered a true alternative to the desktop-based systems.

The replies received to the two questions asked about online tools can be compared to what was presented earlier concerning the implementation of machine translation in TM programs.

Firstly, the amount of user study participants who do not have an opinion on these fairly new kind of translation memory programs is relatively high (11.2% or 66 respondents). This might reflect the fact that many do not know about this technology yet. In total, only slightly more than one third of the respondents

Figure 22. Participants' opinion on online-based translation memory tools

(225 or 38.7%) believe that online tools can truly replace the desktop-based variants that have been established on the market for many years – either judging from their own experience with them at work (114 replies or 19.6%) or from their estimation without actually having used them (111 replies or 19.1%). On the other hand, precisely half of the participants (291) have uttered strong scepticism, believing that online-based programs do not represent a worthwhile alternative. When focusing on those translators whose opinion is based on actual experience with the applications they are meant to judge, the largest share of this group – equalling to around one third of all respondents (195 or 33.5%) – is not satisfied with what these tools offer. Once again, the translation students' replies to this question might raise the hopes of the developers, since only one single student thought that online-based tools do not represent an alternative, with all others being either neutral or positive towards them.

The following (and final) question of the questionnaire, asking the participants whether they currently use an online-based translation memory tool at their workplace, can be seen as the logical consequence of their scepticism and/or partially due to the fact that these tools were only recently made available. Less than one third of the respondents (159 or 27.1%) do actually use one of those tools at the moment, with the rest (428 or 72.9%) not doing so. Developers of this kind of technology might face a struggle against the habits and scepticism of translators towards it, as previous results indicate that it might be a difference between generations.
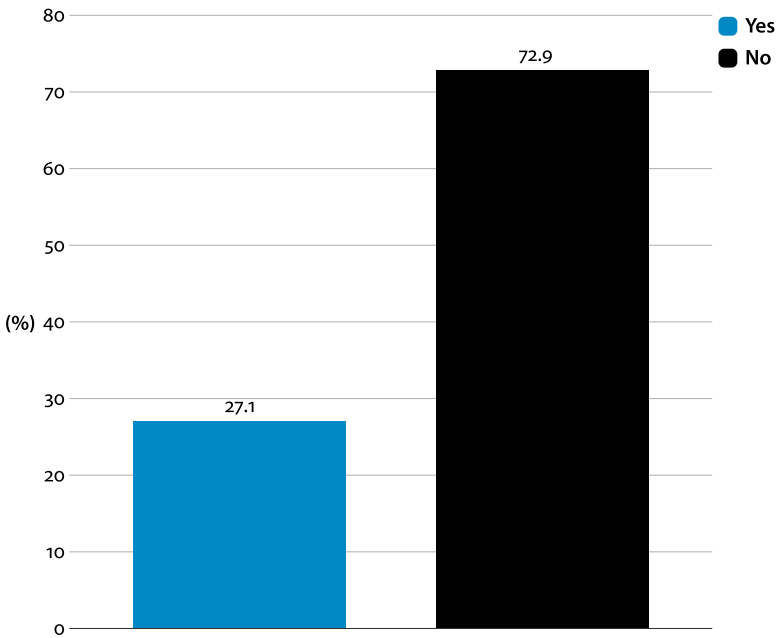
Figure 23. Current workplace use of online-based translation memory tools

4.   Conclusion

This paper presents the findings of a large user survey on translation memory software. As argued in Section 1.1, this survey is the most up-to-date survey on the topic and the one that relied on the broadest audience. Over 700 professionals and students of translation of different backgrounds and specialization have responded the questionnaire.

We presented and analysed the findings in detail in Section 3. Below we provide a short summary of the main findings of our study:

–   Freelance translators have the largest familiarity with TM tools. This might indicate the influence of company- and agency-side enforcement of TM using. Freelancers are typically required to use a particular TM tool if they want to translate to a particular agency.

–   Almost one third of translators are used to translation jobs being bound to specific TM software. In most cases, the requested tool is SDL Trados. Our findings indicate that most users do not carry out extensive comparison of competing software prior to purchase software licenses. This provides less chance for other software to compete because for convenience usually the first choice is the market leader.

- SDL Trados continues to dominate the market by a large margin. However, more than half of its users use at least one other TM program alongside it. This is directly related to the enforcement to use a specific TM software other than Trados.
- Productivity gain, which is evidenced by several studies and claimed by TM software developers, is confirmed by the translators' own perception.
- Project management-based workflow does not seem to meet the translators' requirements. Most respondents regard this as a distraction from the 'actual work'. In our interpretation it is not clear if this refers solely to project management functionalities in CAT tools or to a more widespread negative view on translation project management in general. It should be noted, however, that project management functionalities are a very important feature for language service providers.
- We observed that there is still strong skepticism towards the use of machine translation. This might due to several factors such as previous negative experience with MT, the quality of MT output for certain domains and language pairs or due to the use of non-intuitive post editing software. Even so, as more CAT tools are offering both MT and TM we believe skepticism will decrease in the next years.
- Quality assurance and concordance search are highly important features for users.
- There are a number of translator demands that remain largely unfulfilled and can be addressed by software developers. For example, the need for more context being provided for TM matches.
- Online TM tools are still not considered to be an alternative to traditional TM software. Many users are concerned with privacy and data protection. Skepticism is, however, not as large as towards MT. Even so, in our opinion, this technology has a great potential for growth. Our survey indicates that current workplace use of online TMs is already at 27%.

The comprehensive study presented in this paper provides important information for researchers and developers interested in improving the quality of CAT tools.

Acknowledgements

References

Alabau, Vicent et al. 2014. "CASMACAT: A Computer-assisted Translation Workbench". In *Proceedings of EACL*, 25–28. Gothenburg: Association for Computational Linguistics.

Cettolo, Mauro; Nicola Bertoldi; Marcello Federico; Holger Schwenk; Loïc Barrault; and Christophe Servan. 2014. "Translation Project Adaptation for MT-enhanced computer assisted translation". *Machine Translation* 28 (2): 127–150.
https://doi.org/10.1007/s10590-014-9152-1

Bojar, Ondřej; et al. 2016. "Findings of the 2016 Conference on Machine Translation". In *Proceedings of the First Conference on Machine Translation*, vol. 2, 131–198. Taberg: Association for Computational Linguistics.

Bowker, Lynne. 2005. "Productivity vs Quality? A Pilot Study on the Impact of Translation Memory Systems". In *Localisation Reader*: 133–140.

Gupta, Rohit; Constantin Orăsan; Marcos Zampieri; Mihaela Vela; Josef van Genabith; and Ruslan Mitkov. 2016. "Improving Translation Memory Matching and Retrieval Using Paraphrases". *Machine Translation* 30 (1–2): 19–40.
https://doi.org/10.1007/s10590-016-9180-0

Iarossi, Guiseppe. 2006. *The Power of Survey Design*. Washington, D.C.: The World Bank.
https://doi.org/10.1596/978-0-8213-6392-8

Lagoudaki, Elina. 2006. *Translation Memories Survey 2006 – Translation Memory Systems: Enlightening Users' Perspective.*
http://isg.urv.es/library/papers/TM_Survey_2006.pdf

Federico, Marcello; et al. 2014. "The Matecat Tool". In *Proceedings of COLING (Demos)*, 129–132. Dublin: Dublin City University and Association for Computational Linguistics.

Krosnick, Jon A.; and Stanley Presser. 2010. "Question and Questionnaire Design". *Handbook of Survey Research* 2 (3): 263–314.

LeBlanc, Matthieu. 2013. "Translators on Translation Memory (TM): Results of an ethnographic study in three translations services and agencies". In *Translation & Interpreting* 5 (12): 1–13.

Maguire, Martin; and Nigel Bevan. 2002. "User Requirements Analysis: A Review of Supporting Methods". In *Proceedings of IFIP 17th World Computer Congress*, 133–148. Deventer: Kluwer.

Pal, Santanu; Sudip Kumar Naskar; Marcos Zampieri; Tapas Nayak; and Josef van Genabith. 2016. "CATaLog Online: A Web-based CAT Tool for Distributed Translation with Data Capture for APE and Translation Process Research". In *Proceedings of COLING (Demos)*, 98–102. Osaka: The COLING 2016 Organizing Committee.

Payne, Stanley Le Baron. 1951. *The Art of Asking Questions: Studies in Public Opinion* no. 3. Princeton (NJ): Princeton University Press.

Peterson, Robert A. 2000. *Conducting Effective Questionnaires*. Thousand Oaks (CA): Sage Publications.
https://doi.org/10.4135/9781483349022

Mirko Plitt; and François Masselot. 2010. "A Productivity Test of Statistical Machine Translation Post-editing in a Typical Localisation Context". In *The Prague Bulletin of Mathematical Linguistics* 93: 7–16.
https://doi.org/10.2478/v10108-010-0010-x

Zampieri, Marcos; and Mihaela Vela. 2014. "Quantifying the Influence of MT Output in the Translators performance: A Case Study in Technical Translation". In *Proceedings of the Workshop on Humans and Computer-assisted Translation (HaCat)*. Stroudsburg (PA): Associaton for Computational Linguistics.
https://doi.org/10.3115/v1/W14-0314

Zaretskaya, Anna. 2015. *D2.1: User Requirement Analysis*. EXPERT: EXPloiting Empirical appRoaches to Translation. Technical Report.

Zaretskaya, Anna; Gloria Corpas Pastor; and Miriam Seghiri. 2015. "Translators' Requirements for Translation Technologies: A User Survey". In: *New Horizons in Translation and Interpreting Studies*. pp. 133–134. Geneva: Tradulex.

Résumé

Cet article présente une étude exhaustive sur l'utilisation des mémoires de traduction par des traducteurs d'horizons différents. Nous avons élaboré un questionnaire, qui a été complété par un groupe de 723 répondants, dont des traducteurs professionnels, des étudiants en traduction et des chargés de cours en traductologie et en pratique de la traduction. Nous analysons les résultats de l'enquête en fournissant des informations importantes sur les besoins des utilisateurs, les fonctionnalités les plus importantes d'une mémoire de traduction, la productivité perçue par les utilisateurs et les parts de marché.

Mots-clés: mémoires de traduction, outils d'aide à la traduction, productivité, enquête auprès des utilisateurs

Address for correspondence

Marcos Zampieri
University of Wolverhampton
Wulfruna Street, WV1 1LY
Wolverhampton
United Kingdom
M.Zampieri@wlv.ac.uk

Co-author information

Dominik Schneider
Saarland University
dominikschneider011@yahoo.de

Josef van Genabith
DFKI GmbH and Saarland Unversity
Josef.van_Genabith@dfki.de