Fotini Vallianatou is a freelance translator from English and French into Greek. She has a degree in Translation from the Ionian University, Greece. She has been a freelancer for the past 10 years. She specializes mainly in medical and technical texts.

Ms. Vallianatou can be reached at tsilimb@otenet.gr. Her Web page can be found at www.greek-translations.net.

CAT Tools and Productivity:

Tracking Words and Hours

by Fotini Vallianatou

Abstract

A freelance translator calculates her productivity over a period of 18 months tracking number of words translated, time devoted to translation and the CAT-tool used for every project.

It's been some time since CAT tools have been part of my work routine. When I first started using DejaVu, it was more a curiosity than a real need. At that time CAT tools were a luxury; one had to invest some money, but mainly the translator had to invest time to learn them and to become accustomed with their philosophy. Not everyone was willing to do so. Since then things have changed dramatically. CAT tools are a must today, and for a certain type of translation jobs, there is no chance that you will be busy unless you know how to use them. Moreover, you need to be familiar with more than one commercial package—each client has his own preferences—and sometimes you need to learn a custom-made tool developed by a client for his own needs.

For people like me, who started out in the translation profession in the pre-CAT-tool era and now are regular CAT-tool users, it is obvious that to some extent the main argument of CAT tools developers is true: these devices improve productivity. Yet I always had the curiosity to measure this improvement in terms of time and profit. Especially during the last few years that translation agencies tend to demand considerable price reductions when a translation is processed using a CAT tool, it becomes more and more difficult to understand whether or not the improvement in productivity is canceled by price reductions.

The time to answer this question for my personal practice came when the volume of my work increased to a degree that record-keeping became necessary to track current and older jobs, invoices, clients, files etc. I solved this problem developing my own record-keeping program in Microsoft Access. The program allowed me to keep track of jobs, dates, volumes of words and hours, clients, location of files in my hard disk, invoices, pending jobs and everything else that could make my life easier. It also allowed me to prepare simple statistics, calculating work volume (words or hours or both) for a certain period of time or for a certain client or clients. With the record-keeping program I had a very easy way to calculate words of my previous jobs in any desirable combination.

The simpler way to track productivity with a CAT tool would be to record volume of words translated with this tool together with the time devoted to translation. This would give one a measure of translation output (words/hour). Although such a measurement would not take into account a significant parameter, quality, it could still provide a rough estimate of productivity.

The record-keeping program offered the first precondition necessary for checking CAT-tool-related productivity in my practice. The second would be an easy way to track time while I was translating. The solution to this was the Activity Timer, a shareware program developed by Robert Walker (http://www.tunesmithy.co.uk/). Activity Timer tracks time in an impressively simple way. The last bit was to include into my record-keeping program the option to classify the words of a job the way a CAT tool does (exact match, fuzzy match, no match) and to record which CAT tool I was using every time. With some work in Access, the record-keeping program was able to calculate the time during which I was working using a certain CAT tool and the words that I translated during that time. The rest was based in my commitment to bringing this task to an end.

Today, after 18 months of tracking, I think I know a little more about how CAT tools affect my job, and I would like to share this knowledge with you. In the following paragraphs I will describe in some detail the process of tracking, and I will give you the results of this evaluation.

CAT tools

During the last 2 years, I've been using mainly three CAT tools: DejaVu (latest version X), Trados (latest version 6) and a custom made CAT tool developed by one of my main clients. Thus, this study evaluated these three tools. Translations were processed with one of these tools using either translation memory databases (memories) and terminology databases (terms) provided by the clients or the memories and terms that I had accumulated over the years from my previous projects. I have such memories and terms for both DejaVu and Trados containing approximately 150,000 entries each. Projects processed with the custom-made CAT tool were coming with their own memories, when they were available.

Record-keeping program

The program was developed based on Microsoft Access and currently runs under Windows XP. A form is used for data entry (Fig. 1) and the record is updated whenever a new job arrives. From the various data that are entered for a job, those related to CAT tool/productivity study were:

1. number of words (broken down to no matches, fuzzy matches, and exact matches)
2. CAT tool used (with four selections: Trados, Deja Vu, Other, and None), and
3. total time spent for the job (in hours)

Fig.1: Data entry form of the Record-keeping program. The fields that are used for the CAT tool study are enclosed in blue rectangles.

The number of words is entered as follows: When the client requests a price reduction for fuzzy and exact matches, all three fields are filled accordingly. If the client requests no price reduction, then the total word number is entered in the "no matches" field. This gives me a simple way to group the jobs in "full price" and "reduced price" on the basis of whether or not the fuzzy matches and exact matches fields contain any entries. The request for price reduction is always accompanied by a project-related memory database (memory) from the client. Thus, this grouping also provides a simple way to distinguish portions of jobs that are supported by specialized memories as opposed to jobs that are processed using my own memories.

Activity Timer

This shareware program, developed by Robert Walker (http://www.tunesmithy.co.uk/), tracks time. You can set multiple timers and select an active one each time. One of the main advantages of the program is that you can set a time limit after which the timer stops if the computer is inactive; this way you minimize the possibility that you will accidentally let the timer running when you actually are doing something else. The time limit was set at 1 min. For each translation job that I was tracking, I was creating a new timer and was giving it the same name that the job had in the record-keeping program (copy/paste). After the end of the job, the total time was transferred into the "total time" field of the record-keeping program and the timer was reset.

Calculation of time and words per CAT tool

Calculations are performed per CAT tool. The combination of a number of Access queries permits the calculation of words and hours for the following cases:

1. Trados full price
2. Trados reduced price
3. DejaVu full price
4. DejaVu reduced price
5. Other full price
6. Other reduced price
7. None

A form is used to initiate calculations for a particular period of time (Fig. 2).

Fig. 2: The form that initiates CAT-tool related calculations

The output is a table for each of the cases mentioned above (Fig. 3).

Tr red:no of jobs	Tr red: nm words	Tr red: fm words	Tr red: em words	Tr red:total time
20	38624	14376	72232	113,4

Fig. 3: Table with the results for job processed with Trados for a reduced price. One can see the number of jobs and the totals of "no match" words, "fuzzy match" words, "exact match" words, and translation time.

Criteria for tracking

To be considered for tracking, a job should fulfill the following criteria:

1. It should be a translation job
2. It should not be a "minimum-charge" job, as this kind of jobs were registered without word count, and
3. It should not be recorded as paid by the hour, as some jobs of this type were recorded without counting words.

Results of CAT tool tracking from June 2003 to December 2004

During this period I received 175 jobs, with a total word count of 468,878 words, that fulfilled the tracking criteria. CAT tools and translation times were successfully tracked in 90 of these jobs with a total word count of 359,105 words (76% of the total word load for this period) and a total translation time of 544.85 hours.

The results per CAT tool were as following:

CAT tool	No of projects	Total Word count	Total time
Trados	36	158940	192.8
DejaVu	25	42525	74.35
Custom CAT tool	26	155023	271.2
No CAT tool	3	2617	6.5

Although in the initial design of the study the option of not using a CAT tool was included, at the end of tracking there were only 3 projects in which no CAT tool was used, with a total word

count of 2617 words. Because of the very low word count, this data was considered statistically irrelevant for further processing.

When results were analyzed based on requests for price reduction, the findings were the following:

CAT tool	Price	No of projects	"no match" words	"fuzzy match" words	"exact match" words	Total time
Trados	full	16	33708	0	0	79.4
Trados	reduced	20	38624	14376	72232	113.4

There was only a single DejaVu job done at a reduced price, with a total word count of 1500 words. Because of the low word count, this job was considered not suitable for further analysis. Jobs processed with the custom-made CAT tool had a fixed rate and were not considered in this analysis.

Productivity calculation

Productivity was defined as words translated per hour and it was calculated per CAT tool:

CAT tool	Productivity (words/hour)
Trados (total word count)	824.3
Trados (full price)	424.5
Trados (reduced price)	1104.3
DejaVu	571.9
Custom CAT tool	571.6

Productivity adjusted for price reduction

In order to compare productivity between the two Trados groups (full- and reduced-price) an adjustment was made in the word count of jobs done at a reduced price in order to compensate the income loss caused by the price reduction. The word count of each word category (no match, fuzzy match, and exact match) was multiplied by a factor that incorporated the average percentage of price reduction for each category (1, 0.7 and 0.25, respectively). The word count and productivity were then recalculated.

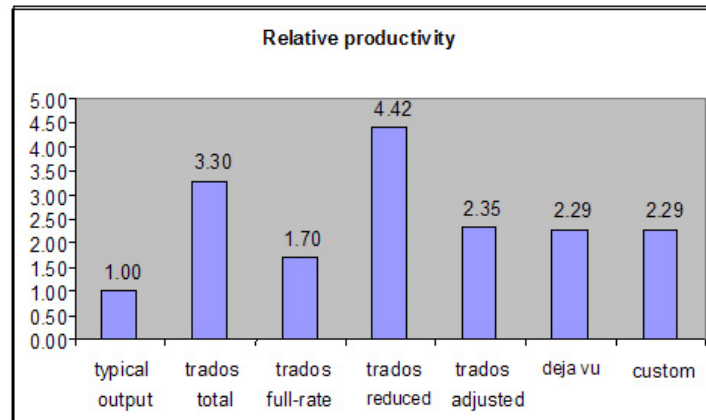| CAT tool	Price	"no match" words x 1	"fuzzy match" words x 0.7	"exact match" words x 0.25	Adjusted word number	Total time (hours)	Adjusted productivity (words/hour)
Trados	reduced	38624 x 1 = 38624	14376 x 0.7 = 10063	72232 x 0.25 =	66745	113.4	588.5

What do all these numbers mean?

Fig. 4: Relative productivity per CAT-tool. The productivity of each tool was compared to the typical output of 250 words/hour which is considered the baseline (relative productivity = 1). (trados total= productivity for the total word count processed with Trados; Trados adjusted=productivity for discounted Trados projects after adjustment of word count to compensate for the price reduction).

The first conclusion from the analysis of these data is that the use of any of the three CAT tools increases productivity considerably. In all three cases productivity was well above the standard average of 250 words/hour. When relative productivity was calculated (defined as the ratio of each individual productivity to the standard output of 250 words/hour), the increase ranged from x 1.6 up to x 4.4 (Figure 4).

Impressive result for all three tools! I had this feeling before I started the study, but seeing the actual results makes any reservations disappear.

Well, how much can one rely on the actual numbers calculated in this study? It is important to remember that these calculations concern my performance at a very specific point of my career (after almost 10 years of freelance work), for a specific type of translation (technical, focused mainly on automotive and medical instrumentation), and for a specific language pair (English into Greek). Moreover, the performance of the CAT-tools is directly related to the content of the actual translation memories used. Therefore, one should be cautious in extrapolating from these data. You can, however, consider them indicative. The typical output of 250 words/hour is—for me at least—a standard that describes well the average output without the use of any kind of CAT tool. The comparison would be much more relevant if I had succeeded to include a considerable volume of words translated without the use of a CAT tool. This was not the case, despite my dedication to this study. I translated only 2617 words (0.7% of the tracked words) without a CAT tool.

What about each individual CAT tool? Although the differences between the three tools are obvious, it is not possible to draw general conclusions concerning the performance of each of them. The actual productivity in each case depends on the type of job and the memory and terminology databases used, even for jobs processed with the same CAT tool. The more specific the memory databases, the larger the word volume you can translate in the same time period. Moreover, each CAT tool may have a different success rate in pulling up existing translations from the translation memories. A detailed evaluation of these parameters would be necessary if one wanted to make comparisons between the tools.

In this study I succeeded to track information concerning matching between the text to be translated and the memories only for a portion of the projects processed with Trados. These projects were supported by specific translation memories supplied by the clients and were priced at a lower rate. There are some very interesting conclusions coming from this group. First, the productivity increase in this group was impressive, reaching almost 5-fold compared to the typical output, underscoring the importance of specialized translation memories. Second, even when the word count was adjusted to compensate for the income loss due to lower rates, the productivity was still almost twice as high as the baseline productivity.

My reluctance to translate without using a CAT tool is by itself an interesting point of this study. At the beginning, I tried to work without a CAT tool for the shake of the study. However, shortly thereafter I decided that this was not acceptable in my everyday professional routine. This is due to at least two reasons. First, the extra time necessary for translation was annoying. Second and most important, in addition to time savings, CAT tools offer invaluable help in terminology and style consistency. This is something that may be difficult to measure, but is

obvious to me, and to any other CAT tool user, I suppose. And finally, since using a CAT tool lets you transfer any project into a memory data base for future reference, I always had the feeling that if I did not use a CAT tool I would waste the chance of a future reference for the same or another client. The option of alignment exists for most CAT tools, of course, but it is time-consuming.

Well, finally at the end of all this effort I confirmed what I already knew empirically. CAT tools are well integrated in my work routine, and they are here to stay! They do speed up the translation process, and they help me provide a better product to my clients. Price reductions, when requested, are indeed annoying, but if the reduction is within acceptable limits, the productivity enhancement is not canceled by the lower rate, and both the client and I are happy. One may have his/her preferences (I have my own) as to which CAT tool is better, but from this study I concluded that all three tools that were tested performed very well and increased my hourly output much above the standard average. I still track my work time. After all these months it comes naturally to turn the Activity Timer on when a new project starts. After all, this habit is not bad: it gives you a better picture of how you allocate your work time...