KNOWLEDGE OF PROVENANCE AND ITS EFFECTS ON TRANSLATION PERFORMANCE IN AN INTEGRATED TM/MT ENVIRONMENT

Carlos S. C. Teixeira
Universitat Rovira i Virgili

A minor dissertation submitted in partial fulfilment of the requirements for the Masters in Translation and Intercultural Studies

Research Masters program in Translation and Intercultural Studies
Universitat Rovira i Virgili, Spain

Supervisor:
Dr Anthony Pym, Universitat Rovira i Virgili

2011

Abstract

The integration of machine translation (MT) and translation-memory (TM) systems in professional translation settings has turned pre-translation + post-editing into an even more attractive alternative in terms of productivity for all parties involved in the translation process. In some cases, source files are pre-translated using a combination of customised MT and TM before reaching the translators, who then become reviewers, or post-editors. But how does this actually affect productivity, and how do translators feel when performing this new activity? In order to look for answers to those questions, we ran a pilot experiment comparing two different environments. The basic difference between the two is the availability of information on the provenance of the suggested translation for a particular segment (whether it comes from MT, TM, and at which match percentage). Data were collected using screen recording, keystroke logging and retrospective interviews. The results show that the differences in overall speed, amount of editing and quality between the two environments are not statistically significant. However, certain types of translation suggestions, namely exact matches and high-percentage fuzzy matches, show pronounced variations in speed and editing.

Keywords: translation technology, translation memory, machine translation, process research, speed, productivity, performance, provenance, trust.

Table of contents

Introduction

Until recently, machine translation (MT) and translation memories (TM) were seen as totally different approaches to using technology in translation. While the first approach was targeted mainly at end users interested in grasping the general idea of a text written in a language they could not understand (usually while browsing the Internet), the second was addressed to professionals in the translation industry, such as translators, translation agencies or translation departments in large companies.

However, this scenario has been changing at a rapid pace in the last few years, mainly due to quality improvements and the general availability of statistical machine-translation systems, based on large amounts of human-produced bilingual text. This has allowed MT to be progressively integrated into TM tools in professional translation environments, bringing new possibilities as well as new challenges.

The potential productivity gains derived from this integration of machine translation (MT) and translation memories (TM) are calling for new work methods in the translation market. As an example, some translation agencies pre-translate their source files using a combination of TM and customised MT before sending them out to translators, who then become reviewers, or post-editors. In this scenario, translators review each segment without knowing its provenance, i.e. whether it comes from a translation memory (and at which match percentage) or from a machine-translation engine. Could this missing information have an impact on the way translators perform their tasks, compared to a more traditional environment, where translators would know where each translation suggestion comes from? In other words, how does the 'knowledge of provenance' of translation suggestions affect translators' behaviour in environments that integrate TM and MT?

Literature review

Machine translation as a technology is much older than translation memories. While the first attempts to create computer programmes that could translate natural languages date back to the 1950s (cf. Hutchins & Somers 1992: 5-6), translation memory systems were conceptualised in the late 1970s, developed for the first time in the mid-1980s and released commercially in the early 1990s (cf. Christensen & Schjoldager 2010: 90).

Due to the technicalities involved in its conception, machine translation has been a subject of study mainly within the computational linguistics community. Translation memory systems, on the other hand, have become a frequent subject within Translation Studies. However, studies on translation memories have been more abundant in the industry – usually with a focus on usability and marketing issues – than in academia (cf. Pym 2011b for the different kinds of contributions that can be expected from either community).

In addition to topics such as productivity and quality, academic research in the field has also addressed the human aspects of technology. Webb (1998) presents a comprehensive study of translation memories systems and summarises the state of affairs at the end of the 1990s. Dragsted (2004) studies the effects of TM systems on cognitive segmentation. O'Brien (2002, 2006) and Guerberof (2002, 2009) study the integration of TM and MT and its effects on productivity, cognitive effort and quality. Krings (2001), Mossop (2001, 2007), Guerra Martínez (2003) and Künzli (2006, 2007) deal with revising and post-editing. Jakobsen (2002) and Dimitrova (2005) study self-revising and compare the performances of experienced and novice translators (cf. Mossop 2007 for an overview of both). Allen (2003, 2005) conducts several studies on machine translation post-editing with specific tools and provides some guidelines for improving its results (cf. also Guerra Martínez 2003). García (2010) compares time and quality between translating "entirely from the source text" vs. "editing machine translation." He finds that "while time differences were not significant, the machine translation seeded passages were more favourably assessed" (2010: 7). Yamada (2011) investigates how the type of content ('free translation' *vs.* 'literal translation') in a translation memory affects translation speed, and concludes that literal translations are more advantageous for higher fuzzy-match categories. Christensen & Schjoldager (2010) provide a good general overview of current research on translation memories and suggest some paths for future research, namely "on how translators interact with TM technology and on how it influences translators' cognitive processes" (op. cit.: 99).

None of the published studies that we are aware of seems to take into account a specific aspect that distinguishes translation memory systems from machine translation systems: TM systems show translators the 'provenance' and the 'quality' of the translation suggestions coming from the memory, whereas MT systems display the 'best translation suggestion possible' without any indication of its degree of confidence. It is our assumption that this missing distinction might be one of the reasons for discrepant results in some studies that compare translation times when (post-)editing MT and TM suggestions. For example, O'Brien (2006) compares the performance of TM *vs.* MT when translators work in a 'traditional' TM system, i.e. where they know the provenance of the translation suggestions they are working with. One of her findings is that "cognitive load [and processing speed] for machine translation matches is close to fuzzy matches of between 80-90% value" (op. cit: 185). For fuzzy matches above 90%, including exact matches, TM processing is faster and requires a lighter cognitive load, whereas the opposite happens for fuzzy matches below 80%. In a different study, Guerberof (2009) reproduces an actual scenario that can be found in industry. The author analyses time and quality when editing translation-memory suggestions *vs.* machine-translation suggestions, in an environment where translators could not tell the provenance of each suggestion. Under this particular condition, her "findings suggest that translators have higher productivity and quality when using machine-translated output than when processing fuzzy matches [at any percentage level] from translation memories" (Guerberof 2009: 11). We can thus say that, at least for high-percentage fuzzy matches, Guerberof and O'Brien have come to rather contradictory conclusions.

Research question and hypotheses

Inspired by the studies by Guerberof and O'Brien briefly mentioned above, we set out to investigate whether the fact of knowing the provenance of the segments could provide and explanation for the discrepancy in their results, especially because no research work in the field seems to have focused on this particular distinction. Our main research question may thus be summarised as: What are the differences (if any) in the translation process between a situation where translators know the provenance of the translation suggestions they are editing and a situation where this information is not available?

In order to answer this question, we compared two translation environments. In the first environment, translators do not know the provenance of translation suggestions, whereas in the second environment translators do have access to this information. These are our working hypotheses:

- Hypothesis 1 (H1): The *translation speed* is higher when *provenance information* is available.
- Hypothesis 2 (H2): The *amount of editing* is smaller when *provenance information* is available.
- Hypothesis 3 (H3): There is no significant difference in the *quality* level when *provenance information* is available.

Some definitions are necessary in order to operationalise the variables we want to test:

- *Translation speed* is measured as words per hour. There are separate counts for the first rendition (drafting) and second rendition (self-revising).

- *Amount of editing* is measured as the percent ratio between the number of keyboard key presses + mouse clicks done by translators and the total number of characters in the resulting segment.

- *Provenance information* of translation suggestions is indicated by showing their origin (TM or MT) and, in the case of TM, by displaying its fuzzy-match percentage and highlighting the differences between the actual segment and the matching segment in the TM, as is usually the case in most TM systems.

- *Quality* is measured as a score given by a professional reviewer, who will process all resulting translations according to predefined criteria (see further details below in the 'Data collection' section).

Methodology

In order to test those hypotheses and fine-tune the methodology for our future doctoral research, we ran a pilot experiment with two translators from English to Spanish. Each of them translated two similar source texts of around 500 words each, in the two different environments described below.

Environment B presents the source-text segments on the left-hand side of the screen and a pre-translated version of the source text (obtained through the pre-processing of the file with TM and MT) on the right-hand side. In this case, all no-matches were replaced with MT suggestions, and the whole text was presented as a sequence of pre-translated segments. Translators could edit the pre-translated segments as if they were revising a translated file and they had no information on the origin of each of the pre-translated segments (i.e. whether they came from a TM segment or an MT segment). For mnemonic recall, let us call this environment B (as in 'Blind'). This environment tries to reproduce most of the conditions found in the environment described in Guerberof (2009).[i]

Environment V is similar to the previous one in that translators also had access to the source-text segments on the left-hand side of the screen and an editing space on the right-hand side. However, its difference consists in that, instead of working by 'blindly' editing pre-translated segments, translators could see where the default translation suggestion was coming from (either from the translation memory or from the MT engine). Additionally, in the case of TM suggestions, translators could see the highlighted differences between their actual source segment and the TM source segment. For mnemonic recall, let us call this environment V (as in 'Visual'). This environment tries to reproduce as close as possible the environment described in O'Brien (2006).

Data collection

The main methods for collecting data were screen recording and keystroke logging through BB FlashBack Express 2. We also considered using Translog for the latter task, but since it has its own interface and does not allow for integration with other applications, it would not be possible to reproduce the two environments we wanted to test (cf. Jakobsen 2002, 2003 on the use of Translog for keystroke logging). The two translation environments were created within SDL Trados Studio 2009 Freelance.

Think-aloud protocols were not used as they are known to slow down the translation process (cf. Krings 2001, Jakobsen 2003) and we were precisely trying to measure translation speed in a natural(istic) environment. Retrospective interviews can provide the information we need on translators' feelings and satisfaction, without interfering with the translation process. Eye tracking is an additional data-collection method that is being considered (cf. O'Brien 2006), pending on the availability of equipment, and would provide a more precise way of identifying which segment is being focused on at a given point in time.

For assessing quality, all texts were rated by two reviewers, based on an error-count system. The quality level of a translation was measured through a score, which starts at 10 and decreases according to the grid shown in Table 1.

Table 1. Quality assessment grid

Type of error	Deduction
1 misspelled word	.25
1 grammar mistake (morphology, syntax)	.25
1 use of incorrect or inconsistent terminology	.25
1 general readability (understanding) issue	.25
1 sentence structuring issue (style, register)	.25
1 instance of omitted information	.25
1 instance of incorrect or inaccurate meaning rendition	.25
1 localisation error (numerical formats, units)	.25
Other deductions	.25 each

Data processing

Speed was measured by dividing the total time translators used while translating a segment by the total number of words in the source segment. BB FlashBack shows the translators' face and screen activity at all times and helps us identify whether the translator is looking at the source text (left half of the screen) or at the target text (right half of the screen), but it does not allow us to know precisely at which segment translators are looking at a specific moment. For this, we need to rely on keyboard and mouse activity and there is some guesswork involved (cf. the Discussion section). As mentioned above, an eye-tracker would be useful for solving this problem.

The amount of editing was measured as the ratio between the number of keyboard keys pressed + mouse clicks and the total number of characters in the resulting segment. For example, if a translator cuts the last half of a 36-character segment by pressing Ctrl-X (1 mouse click + 2 key presses), then moves the mouse to the beginning of the sentence and clicks (1 click), pastes the text at the beginning by pressing Ctrl-V (2 keys), then types 3 characters, then presses Delete (1 key), then moves and clicks the mouse again (1 click), then finally presses Delete (1 key) and types 1 character, the amount of editing for this segment is (3+1+2+3+1+1+1+1)/36 = 36%. The pressing of Shift for producing uppercase letters or the pressing of a key combination to produce an accented character counts as a single key press. There might be cases in which the amount of editing is greater than 100%, meaning that the translator pressed more keys than the total number of characters in the final segment. This is an arbitrary measurement system, but since it is a consistent way of measuring the same variable in two situations, it shall not affect the validity of our results.

Some special cases to be considered. If the translator uses the arrow keys just to move around the text as a reading aid, those key presses are not counted. If the translator places the cursor at the position they want to edit, either by using the arrow keys several times or by moving the mouse and clicking once, this movement counts as one keystroke. If the translator selects text for copying or cutting, by using either the mouse or key combinations, this action counts as two keystrokes. Edits made during automatic spell checking are counted as one keystroke if the ‘Change’ button is used, or are counted as usual if the spell-check window is closed and the text is corrected by typing on the target segment.

Appendices 10-13 show the tables used to keep track of the times spent and the edits performed in each of the segments by each of the subjects for each of the environments.

Pilot experiment

Subjects

Both subjects are male and native speakers of Spanish. Subject1 has formal training in translation and around 4 years of professional experience in several fields, especially audiovisual translation. Subject2 also has formal training in translation and around 8 years of professional experience in various fields, mainly in localisation and technical translation. Both are familiar with many different translation memory systems. This information was obtained through a pre-assessment questionnaire answered by email (see Appendix 1).

Materials

Our source texts were taken from an article in a magazine called *Metalworking World*, published by Spoon Publishing, Sweden. The publication is subtitled “a business and technology magazine from Sandvik Coromant,” a Swedish metalcutting tool manufacturer, and is distributed worldwide to the company’s customers. A master version is written in English and is then sent out for translation in order to have the magazine localised for 19 different languages/markets. The source texts used in the experiment deal with composite materials in car manufacturing and were taken from an article in issue 2.11 of the magazine. The main reason for choosing this kind of material was a wish to use technical text outside of the ‘software localisation’ domain – which is the object of most research studies in the field – still with (marketing) stylistic features that make it more demanding for translation. The specific article was chosen mainly because of its topic (which should be somehow interesting for translators) and length – allowing for the extraction of two excerpts of around 500 words.

The decisions to do intra-subject comparisons, use authentic source texts that were as similar as possible to each other, give translators realistic instructions, and have translators work in a setting as close as possible to their natural working environment find support in Pavlović (2007: 60-1):

> Livbjerg and Mees (2003: 125) […] discuss some problems related to research design with regard to the choice of texts: it is not possible to have the same subjects translate the same (unknown) text in two different conditions. An alternative is to use different texts that are judged to be about the same degree of difficulty, although this is not all that easy. Some authors (e.g. Neunzig 2001: 88) propose using different sections of the same text. Another way out of the predicament is to use the same text with different but comparable subjects. But according to Laukkanen (1996: 265), using the same subjects means that “the effect of the subject’s personality is the same in both tasks which is a clear advantage compared to experiments where different subjects’ processes are compared”. […] Most authors agree, however, that authentic texts should be used in the experiments, and that the subjects should be given realistic task descriptions (a “brief”). (Pavlović 2007: 60-1)

Text and memory preparation

The chosen article had a total of 1310 words, corresponding to 55 source segments, with 23.8 words per segment in average. In order to have two source texts of around 500 words, it was necessary to use 21 segments for each of them. As a result, SourceText1 has 512 words (see Appendix 4) and SourceText2 has 510 words (see Appendix 5).

A translation memory was created by aligning the English source text with the Spanish target text (final Adobe InDesign version revised by a copy-editor and approved by the client) using SDL Trados WinAlign + manual verification of each segment. Formatting tags were removed on both sides, as they were very inconsistent between source and target and the visual presentation of the article was not relevant for the experiment.

When the segmentation was different between the English source and the Spanish target segments, the following criteria were used for alignment: (a) when one source segment corresponded to more than one target segment, the target segments were joined into one segment; (b) when more than one source segment corresponded to one target segment, the target segment was split into the same number of segments as the source segments, in the most appropriate location as far as meaning was concerned.

Considering the total number of segments available (21) and the characteristics of each type of match, a decision was made to use the following fuzzy match distribution in the experiment:

- 7 ‘no matches’ (replaced by MT feeds);
- 5 exact matches;
- 9 fuzzy matches, of which
    - 3 matches within the 70%-79% range,
    - 3 matches within the 80%-89% range, and
    - 3 matches within the 90%-99% range.

The order of presentation of match types during translation was defined by a random number generator and it was different for each of the environments. To obtain two memories with the characteristics above, one for each environment, first we converted the exported aligned memory into a table format, so it was visually easier to work with. Next, we selected the segments corresponding to each of the source texts and split the main memory in two, discarding the extra segments that did not correspond to either of the source texts. Then in each of the memories we deleted 7 matches in the order previously defined for the ‘no matches’, left the 5 exact matches untouched, and edited the remaining 9 fuzzy matches to obtain the desired percentage. In order to obtain those fuzzy matches, we resorted to the following strategies: delete parts of the source and target segments, include or replace some words in the source and target segments, or edit the source text. Finally, ‘no matches’ were replaced through SDL Trados Studio with translation suggestions provided by the public, freely available Google Translate machine-translation engine. The two resulting translation memories are presented in Appendices 6 and 7.

The last step was to create two separate projects in SDL Trados Studio: the first project had Text1 as the source text and emulated environment B; the second project had Text2 as the source text and emulated environment V. In order to emulate environment B, we edited the translation memory to have all segments display as 100% matches, pre-translated the file and removed the translation memory from the project.

Running the experiment

Each translator was tested on a different day in similar classrooms at the URV. Both subjects chose to use their own laptop computers during the experiment. Before they started, we made sure they had the required version of SDL Trados and BB FlashBack installed and configured. The aim was to have translators work in an environment as close as possible to their natural work environment. In practice, working with their own equipment meant they could keep their preferred configuration in terms of keyboard, screen and mouse (either built-in or external), operating system (within the Microsoft Windows family), browser favourites, dictionaries, etc. They also had access to the Internet during the experiment.

At the beginning of the experiment, a digital voice recorder was turned on. The initial tasks subjects had to perform were: (a) copy a short passage (see Appendix 2) in Spanish, and (b) translate a short passage (see Appendix 3) from English to Spanish. In both tasks, the source texts were printed out and translators had to type their target texts in Microsoft Word. The purpose of these two initial tasks was twofold: to measure their baseline typing speed (to assess whether this had an influence on their editing strategies) and to serve as a warm-up (and stress-down) activity. This came from a suspicion that the typing ability of each individual translator might have an influence on their performance with each kind of translation suggestion.

Next, the translators were given instructions in Spanish on how to perform the main tasks for the experiment.[ii] Appendix 8 shows the instructions for environment V and Appendix 9 shows the instructions for environment B. The sentences in bold indicate the main differences between the two environments. In general terms, the instructions told the subjects that the memory they were using was created based on a client-approved final version of the Spanish magazine, that it contained different kinds of matches, and that machine translation was used for the remaining segments. Since the translation 'brief' mentioned they would be paid the same amount per word (no fuzzy-match discounts), it was implied that translators were supposed to revise all segments, including exact matches. The instructions also made it clear that their translations were going to be assessed and graded for quality by a professional reviewer. Thus, the instructions implicitly asked translators to try to achieve maximum quality in both environments. Finally, a time limit of 1.5 hours was set for each of the texts.

There was an optional preparation phase to make sure translators knew how to use the translation tool. Subject1 chose to translate some segments of this preparatory text, while Subject2 did not. At this moment, BB FlashBack was turned on to start recording the translators' performances. The tool was set to record screen activity; keystrokes; mouse position, movements and clicks; translators' faces; and sound (voices, keyboard, etc).

Before they started to work on the first main text, subjects were provided with a printed copy of the source text, which was placed on the table or stuck on a spare computer monitor beside their working computers. This later helped us see when they were looking at the printed source text or at the main computer screen. Subject1 spent 2 min 14 sec reading SourceText2 before starting to translate it in the V environment and apparently did not read SourceText1 before starting to translate it in the B environment. Subject2 spent 2 min 55 sec reading SourceText1 before starting to translate it in the B environment and spent 2 min 34 sec reading SourceText2 before starting to translate it in the V environment. In general, neither of the subjects consulted the printed source text again after starting the actual translation.

Results

Time was measured by watching each of the translators' performances in BB FlashBack Player and manually noting down the start and end times for each individual task (see Appendices 10-13). Time was counted when translators were typing, thinking, hesitating, or looking at the source text (except when they read the full source text before starting the translation, as we cannot make a correspondence between the time and specific segments). Time was not counted when translators switched to another window to look up terminology, tried to find a specific function in the tool, or spoke with the researcher. Time spent while thinking of a term already accounts for additional time related to term lookup and works as a kind of compensation, eliminating excess overhead time related to searches in external sources. The time counter was paused when the subject started moving the mouse to go to another application (usually a web browser) outside of the translation environment. It was also paused when the subject moved to the source segment to copy text to be pasted in the browser. Time count was resumed when subject returned to the translation environment (after two frames = 140 msec). Time spent on searches within the translation environment (mainly with the Concordance function) was considered as translation time.

The tables in appendices 14-17 show the data on speed and amount of editing consolidated per type of suggestion for each subject in each of the environments. The tables show separate results for the drafting phase (first rendition) and for the self-revising phase (second rendition), and also the combined results for both phases.

Speed – Subject1

Table 2 shows the average speed results for Subject1.

Table 2. Average translation speed per type of segment in each of the environments for Subject1.

		SOURCE WORDS	TIME (sec) 1st rendition	SPEED (words/hr) 1st rend.	TIME (sec) 2nd rendition	SPEED (words/hr) Combined
'VISUAL' ENVIRONMENT (V)	EXACT (100%) MATCHES	131	155	3036	94	1895
	90-99% MATCHES	91	234	1397	101	977
	80-89% MATCHES	51	153	1197	27	1019
	70-79% MATCHES	87	401	781	88	577
	NO MATCHES (MT FEEDS)	150	783	690	132	591
		510	**1727**	**1063**	**441**	**847**
'BLIND' ENVIRONMENT (B)	EXACT (100%) MATCHES	128	566	815	49	749
	90-99% MATCHES	65	369	635	53	555
	80-89% MATCHES	77	210	1321	30	1155
	70-79% MATCHES	77	273	1014	20	946
	NO MATCHES (MT FEEDS)	165	592	1004	129	825
		512	**2009**	**917**	**281**	**805**

From Table 2, if we look at the average results for the first rendition (drafting), we see that translation speed is higher in V (1063 words/hr) than in B (917 words/hr), a difference of 15.9 percent. If we look at the results for the first and second renditions (drafting + self-revising) combined, translation speed is still slightly higher in V (847 words/hr) than in B (805 words/hr), but the difference is reduced to 5.2 percent. Due to the dispersion of the data (see appendices 14 and 15) and the reduced number of segments in the texts, this variation has no statistical significance.

However, if we look at the different fuzzy-match levels, differences in speed start to stand out. Figure 1 below shows average translation speeds according to fuzzy-match levels for the ‘visual’ (V) environment.

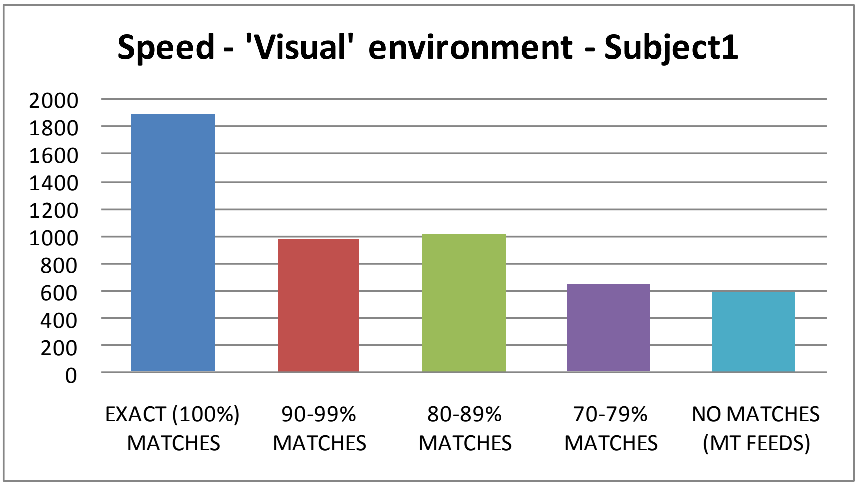

Fig. 1. Combined average translation speed (words/hr) according to fuzzy-match levels for the 'visual' (V) environment – Subject1.

From this bar chart, we can identify three groups of speed levels: exact matches are translated the fastest, fuzzy matches between 80-99% are translated at around half that speed, and lower fuzzy matches (below 80%) and MT output are translated the slowest. This is in accordance with intuitive expectation and with the results obtained by O'Brien (2006).

Figure 2 shows average translation speeds according to fuzzy-match levels for the 'blind' (B) environment.

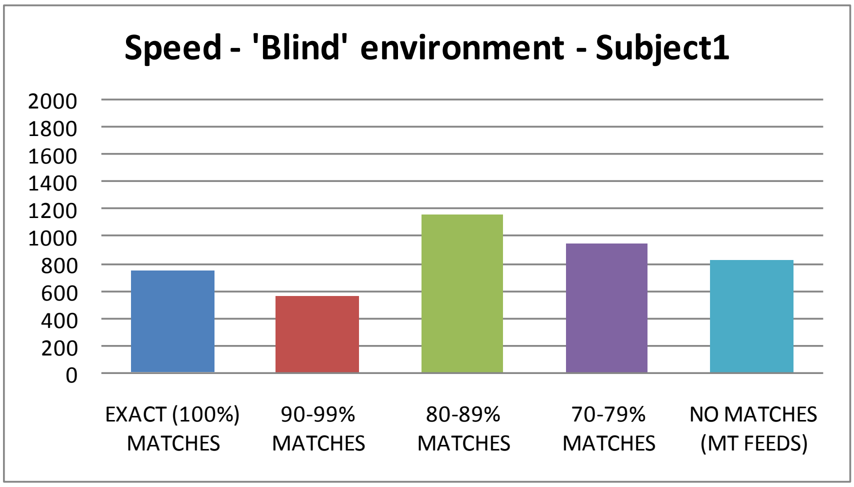

Fig. 2. Combined average translation speed (words/hr) according to fuzzy-match levels for the 'blind' (B) environment – Subject1.

If we look at the exact matches, we notice a dramatic reduction in speed (from 1895 to 749 words/hr), suggesting that provenance information has a high impact on this kind of translation suggestion. Matches in the 90-99% range also show a dramatic reduction in speed (from 977 to 555 words/hr), again indicating that provenance information has a significant impact in this case. Matches in the 80-89% range did not show a significant variation. For lower fuzzy matches and MT feeds, it is worth noting that there was an *increase* in speed.

Figure 3 below shows a comparison of average translation speeds for different kinds of translation suggestions between the two environments.

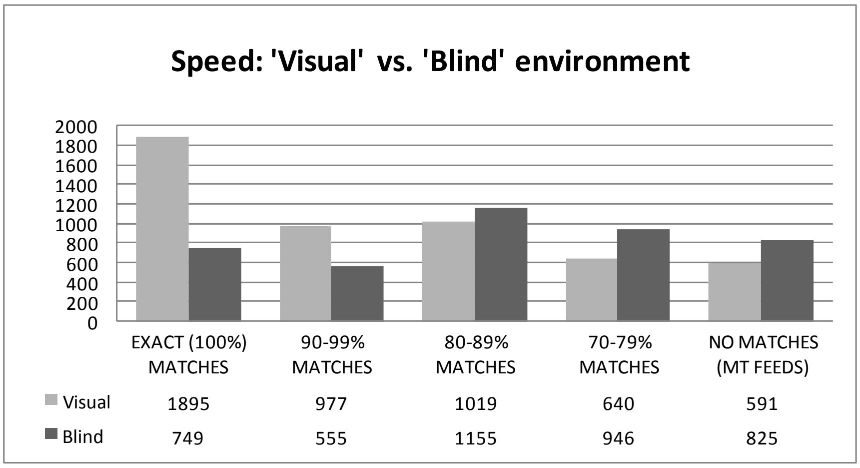

Fig. 3. Combined average translation speed (words/hr) according to fuzzy-match levels comparing environments V and B – Subject1.

Speed – Subject2

Table 3 shows the average speed results for Subject2.

Table 3. Average translation speed per type of segment in each of the environments for Subject2.

		SOURCE WORDS	**TIME (sec) 1st rendition**	**SPEED (words/hr) 1st rend.**	**TIME (sec) 2nd rendition**	**SPEED (words/hr) Combined**
'VISUAL' ENVIRONMENT (V)	EXACT (100%) MATCHES	131	236	2000	121	1323
	90-99% MATCHES	91	354	925	160	637
	80-89% MATCHES	51	225	814	77	606
	70-79% MATCHES	87	456	687	135	530
	NO MATCHES (MT FEEDS)	150	475	1138	214	784
		510	1746	1052	708	748
'BLIND' ENVIRONMENT (B)	EXACT (100%) MATCHES	128	445	1035	95	854
	90-99% MATCHES	65	275	852	52	716
	80-89% MATCHES	77	226	1229	143	752
	70-79% MATCHES	77	289	961	79	755
	NO MATCHES (MT FEEDS)	165	568	1045	161	814
		512	1802	1023	530	790

For this translator, the average results for the first rendition (drafting) show that translation speed is also higher in V (1052 words/hr) than in B (1023 words/hr), but the difference is much smaller than for Subject1, at only 2.8 percent. The combined results for the first and second renditions (drafting + self-revising) show that translation speed is now higher in B (790 words/hr) than in V (748 words/hr), with a difference of 5.6 percent. As was the case with the data for Subject1, this difference is not statistically significant.

Now let us look again at the speed differences according to the various fuzzy-match levels. Figure 4 below shows average translation speeds according to fuzzy-match levels for Subject2 in the ‘visual’ (V) environment.

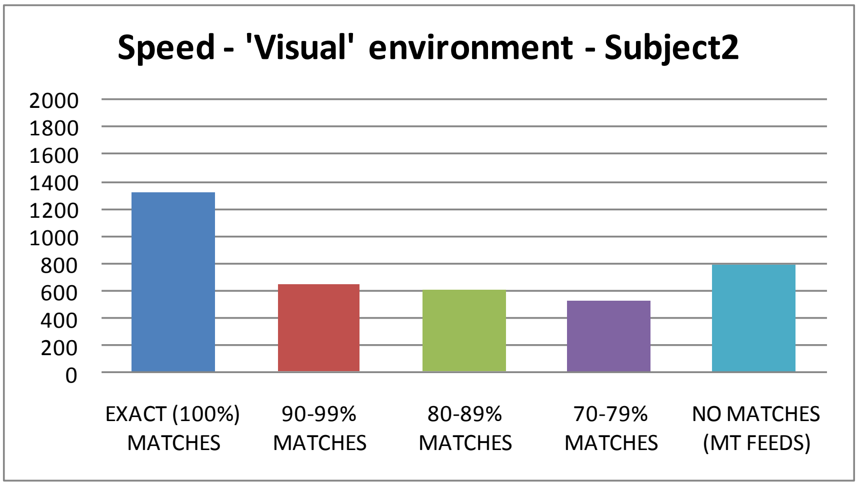

Fig. 4. Combined average translation speed (words/hr) according to fuzzy-match levels for the 'visual' (V) environment – Subject2.

Roughly speaking, this bar chart shows that Subject2 processed translation suggestions coming from exact matches two times faster than suggestions coming from fuzzy matches (1323 *vs.* 591 words/hr in average), and he translated suggestions coming from machine translation around 33 percent faster than the average speed for fuzzy matches. The faster speed for exact matches is still in accordance with our expectations, but the reasons for machine-translation suggestions being translated faster than high-percentage fuzzy matches should be investigated further.

Figure 5 shows the average translation speeds according to fuzzy-match levels for Subject2 working in the 'blind' (B) environment.

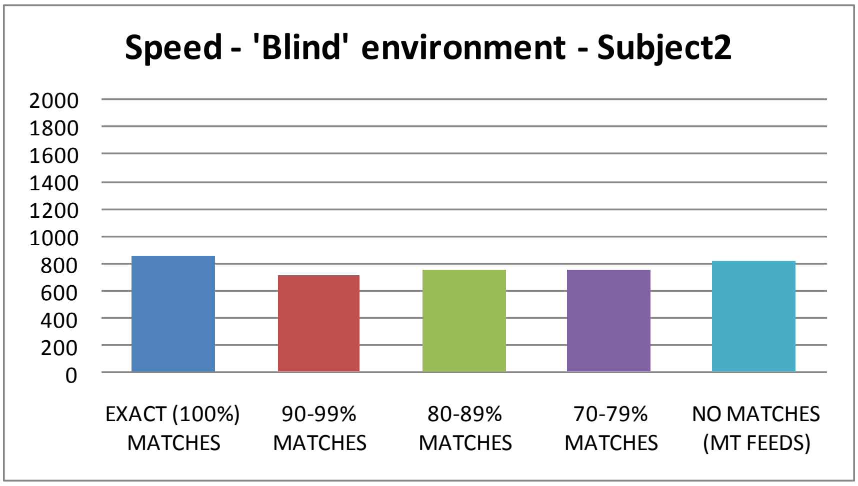

Fig. 5. Combined average translation speed (words/hr) according to fuzzy-match levels for the 'blind' (B) environment – Subject2.

Similarly to what happened with Subject1, the data for Subject2 indicate a dramatic reduction in the average translation speed (from 1323 to 854 words/hr) for suggestions coming from TM exact matches. All other kinds of translation suggestions had an *increase* in speed, with fuzzy matches in the 80-89% range showing the largest increase (42.5 percent). It is interesting to note that differences in translation speeds tend to disappear in the blind environment: exact matches were translated slightly faster, at 854 words/hr, followed by machine-translation suggestions, at 814 words/hr, with translation-memory fuzzy matches being translated a little more slowly, between 716 and 755 words/hr. If the statistical errors are taken into account, differences between the five types of translation suggestions are actually not significant. Figure 6 shows a comparison between average translation speeds for different kinds of translation suggestions in the two translation environments.

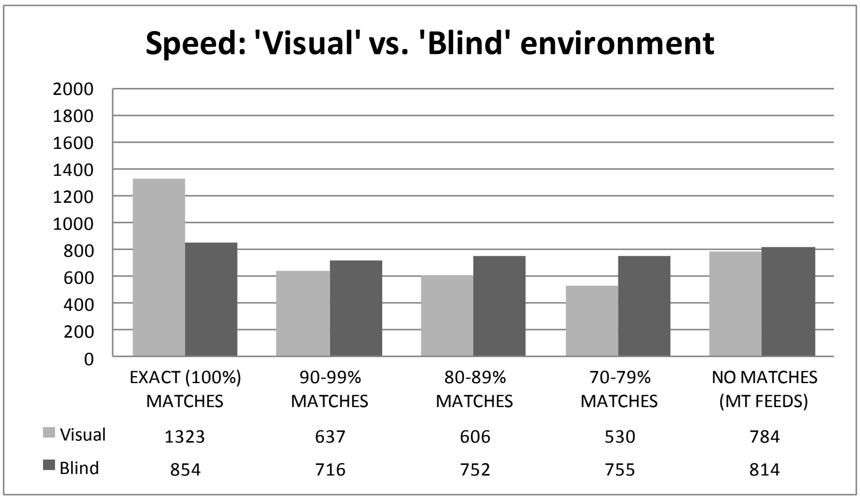

Fig. 6. Combined average translation speed (words/hr) according to fuzzy-match levels comparing environments V and B – Subject2.

Amount of editing – Subject1

For each translated segment, we also computed the amount of editing, which was calculated by taking the number of keystrokes plus mouse clicks used by the translator and dividing this number by the total number of characters in the final target segment (cf. the Methodology section for a more detailed explanation on this calculation). Table 4 shows the corresponding average results for Subject1.

Table 4. Average amount of editing per type of segment in each of the environments for Subject1.

		TARGET CHARS	TYPED CHARS 1st rendition	AMOUNT OF EDITING 1st rendition	TYPED CHARS 2nd rend	AMOUNT OF EDITING Combined
'VISUAL' ENVIRONMENT (V)	EXACT (100%) MATCHES	792	0	0%	11	1.4%
	90-99% MATCHES	464	30	6.5%	75	22.6%
	80-89% MATCHES	376	145	38.6%	0	38.6%
	70-79% MATCHES	457	160	35.0%	0	35.0%
	NO MATCHES (MT FEEDS)	1018	401	39.4%	3	39.7%
		3107	736	23.7%	89	26.6%
'BLIND' ENVIRONMENT (B)	EXACT (100%) MATCHES	792	302	38.1%	9	39.3%
	90-99% MATCHES	468	209	44.7%	23	49.6%
	80-89% MATCHES	416	161	38.7%	5	39.9%
	70-79% MATCHES	418	180	43.1%	0	43.1%
	NO MATCHES (MT FEEDS)	1069	306	28.6%	53	33.6%
		3163	1158	36.6%	90	39.5%

In general, for Subject1 we can say that great amounts of editing in the first rendition implied smaller amounts of editing in the second rendition. This is in accordance with our general expectations, but it highlights the importance of using the combined amount of editing for measuring the total typing effort for each segment. Figure 7 below shows the combined average results for the visual environment.

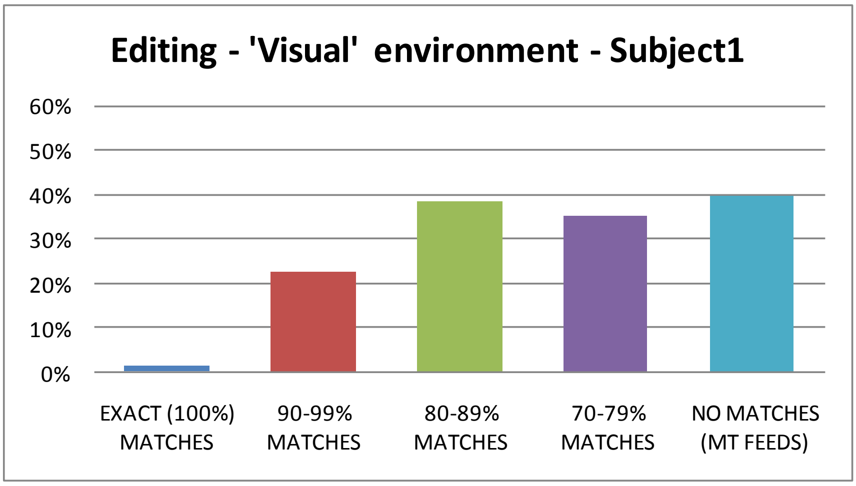

Fig. 7. Combined average amount of editing for each type of translation suggestion in the 'visual' (V) environment – Subject1.

There is a clear difference between the amount of editing (AE) for exact matches (only 1.4 percent) and the AE for the remaining types of suggestions. Fuzzy matches in the 90-99% range required 22.6 percent of the total editing effort, while lower matches and MT feeds required the most effort (between 35 and 39.7 percent). Although fuzzy matches in the 80-89% range required a significantly greater amount of editing than fuzzy matches in the 90-99% range, it is worth recalling that both types of translation suggestions were processed at approximately the same speed (see Table 2), which might indicate that the two variables tested in H1 and H2 are not necessarily correlated.

As for the amount of editing in the 'blind' environment, we notice an increase in all types of suggestions, except for machine translation. However, the difference between the various types of translation suggestions is not as pronounced as in the 'visual' environment, as can be seen in Figure 8.

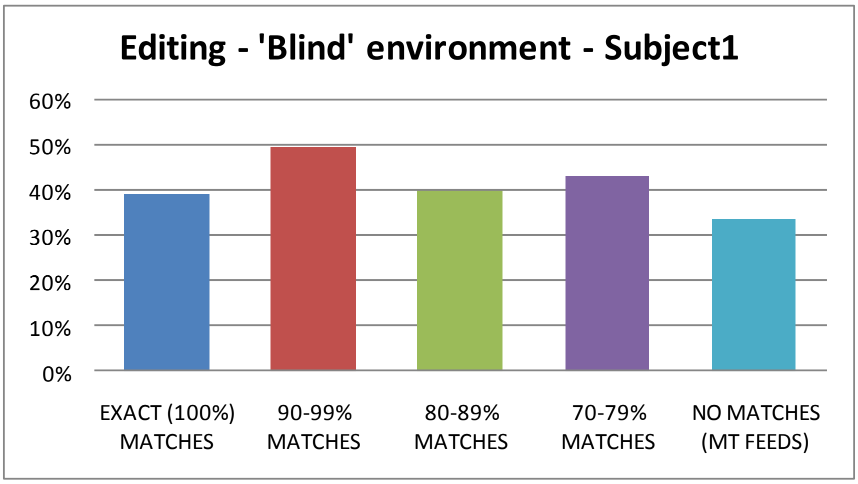

Fig. 8. Combined average amount of editing for each type of translation suggestion in the ‘blind’ (B) environment – Subject1.

The variation in the AE between the two environments is greatest for exact matches (from 1.4 to 39.3 percent) and for 90-99% fuzzy matches (from 22.6 to 49.6 percent), as can be seen in Figure 9.

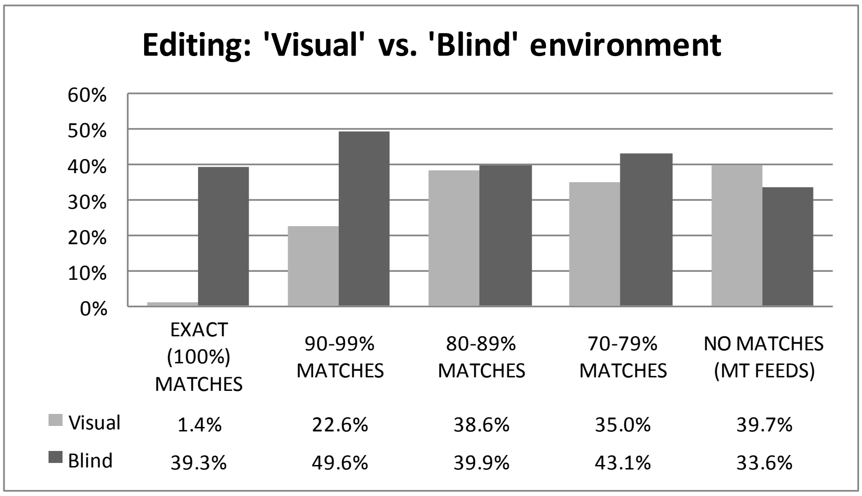

Fig. 9. Combined amount of editing for each type of translation suggestion comparing environments V and B – Subject1

Amount of editing – Subject2

Table 5 shows the average amount of editing per type of translation suggestion for Subject2.

Table 5. Average amount of editing per type of segment in each of the environments for Subject2.

		TARGET CHARS	TYPED CHARS 1st rendition	AMOUNT OF EDITING 1st rendition	TYPED CHARS 2nd rend	AMOUNT OF EDITING Combined
'VISUAL' ENVIRONMENT (V)	EXACT (100%) MATCHES	799	36	4.5%	0	4.5%
	90-99% MATCHES	492	319	64.8%	64	77.8%
	80-89% MATCHES	321	93	29.0%	23	36.1%
	70-79% MATCHES	479	225	47.0%	35	54.3%
	NO MATCHES (MT FEEDS)	984	366	37.2%	72	44.5%
		3075	1039	33.8%	194	40.1%
'BLIND' ENVIRONMENT (B)	EXACT (100%) MATCHES	818	291	35.6%	17	37.6%
	90-99% MATCHES	483	196	40.6%	3	41.2%
	80-89% MATCHES	441	275	62.4%	155	97.5%
	70-79% MATCHES	457	223	48.8%	31	55.6%
	NO MATCHES (MT FEEDS)	1071	325	30.4%	54	35.4%
		3270	1310	40.0%	260	48.0%

The mean results for Subject2 indicate a greater overall amount of editing than those for Subject1 in both environments (40.1 *vs.* 26.6 percent in V and 48 *vs.* 39.5 percent in B) and for most types of translation suggestions. On the other hand, Subject2's translation speeds were only slightly lower than those for Subject1 (748 *vs.* 847 in V and 790 *vs.* 805 in B), reinforcing the idea that speed and amount of editing might not be strongly correlated, as was expected from H1 and H2.

If we look more closely at the 'visual' environment, we see that exact matches required the least typing effort (4.5 percent) from Subject2, as was also the case for Subject1. Then we see an unexpectedly high amount of editing for suggestions in the 90-99% fuzzy range. This is probably due to a single segment (see Appendix 16), for which the amount of editing was extremely high (122.6 percent), although the data for two other kinds of translation suggestions also had large variance. Figure 10 below shows the average amount of editing in the visual environment.

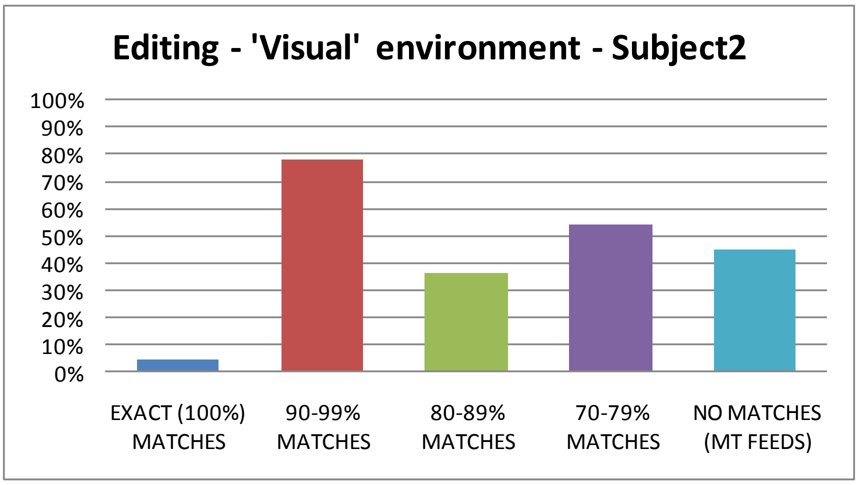

Fig. 10. Combined average amount of editing for each type of translation suggestion in the 'visual' (V) environment – Subject2

Results for the 'blind' environment show that suggestions within the 80-89% fuzzy range required the most editing effort. This can again be explained by an exceptionally great amount of editing (167.4 percent) in a specific segment, in which the translator virtually rewrote the whole translation in the first rendition and then again in the second rendition. As for the remaining suggestion types, their amount of editing lies between 35 and 41 percent, except for fuzzy matches within the 70-79% range. Figure 11 illustrates these data in a clearer way.

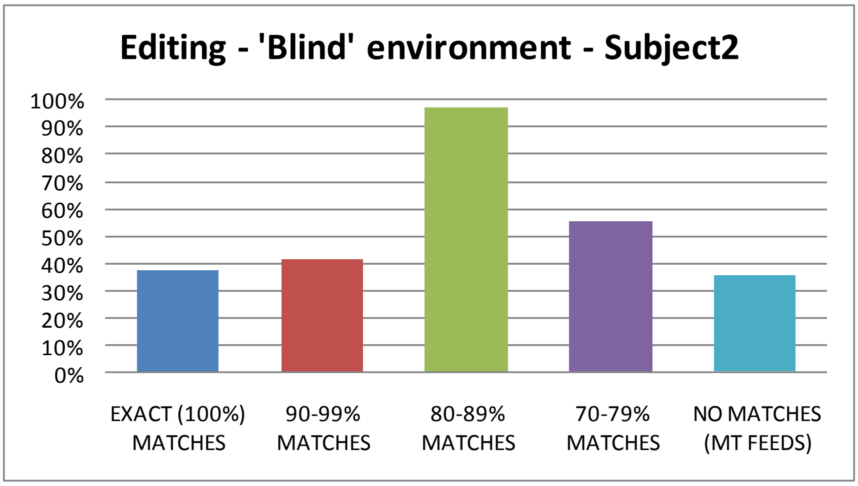

Fig. 11. Combined average amount of editing for each type of translation suggestion in the 'blind' (B) environment – Subject2.

The difference in the AE between the two environments is significant for exact matches (increase from 4.5 to 37.7 percent), for 90-99% fuzzy matches (decrease from

77.8 to 41.2 percent) and for 80-89% fuzzy matches (increase from 36.1 to 97.5 percent), as can be seen from Figure 12.

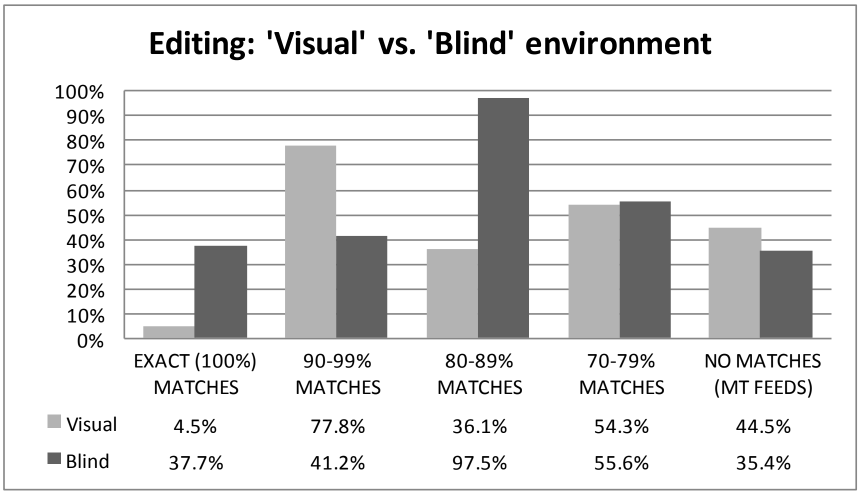

Fig. 12. Combined average amount of editing for each type of translation suggestion comparing environments V and B – Subject2.

Quality

Two revisers assessed the quality of the four translations using the grid provided in the Methodology section above. Revisers were then told to compare the two translations from the same subject and decide which one was better, if any, and to give their final grade from 0 (worst) to 10 (best). The results are shown in Table 6.

Table 6. Translation quality levels for both subjects.

	Subject1		*Subject2*	
	Text 1 (environment V)	Text 2 (environment B)	Text 1 (environment V)	Text 2 (environment B)
Reviser 1	8.5	7.0	8.5	9.0
Reviser 2	7.5	7.0	8.0	8.5
Average	***8.0***	***7.0***	***8.25***	***8.75***

According to the two evaluators, Subject1 performed better in environment V, while Subject2 performed slightly better in environment B. From the evaluators' feedback, we think that quality assessment has not been done properly and the above grades need to be revised again before we can make any definite conclusions. Furthermore, we think the rating instructions need to be made clearer and a greater number of revisers shall be used.

Discussion

We took pains to control most of the factors that might affect our results (type of text, length of text, source language, target language, translator's experience, translation tool, etc.) and we tried to have only our main independent variable (knowledge of provenance) act on our three dependent variables (speed, amount of editing and quality). However, we are aware that many potential extraneous variables (confounds) were also present and had not been properly considered.

Although our aim was to be able to draw some conclusions from a pair of intra-subject studies, the translators' personal styles played a more prominent role than we had originally expected. A more subjective analysis made by the researcher while watching the video recordings and listening to the interviews might give some hints. For example, Subject1 might have produced slightly lower quality translations because he 'respected' the suggestions from the translation memory more often than Subject2, and one of the evaluators did not like the solutions present in the memory.

Data from our pilot experiment do not allow us to draw a definite conclusion on our first hypothesis (on speed) if we take the whole texts as a reference. Subject1 was slightly faster (5.2 percent) in environment V, while Subject2 was slightly faster (5.6 percent) in environment B. However, we can assume that the overall speed, besides individual-specific differences, depends on the distribution of different types of translation suggestions in the texts, as both subjects were faster with certain types of suggestions. For example, if our texts contained only exact matches and machine translation feeds, our results for the entire texts would be different.

As far as our second hypothesis is concerned, our data seem to support it, as the amount of editing for Subject1 was 48.6 percent higher in environment B than in environment V, and the amount of editing for Subject2 was 19.7 percent higher in environment B than in environment V. It is interesting to note, however, that there seems to be no correlation between translation speed and amount of editing, which goes against our expectations.

Our third hypothesis needs further verification, as the data on quality we have for each subject are not sufficient to determine whether the translations produced in the two environments are of the same quality for each individual subject or not.

Data from the retrospective interviews have helped shed light on the translators' general impressions in each environment, as well as on their feelings as 'post-editors'.

Limitations

One advantage of running a pilot test is that it allows for testing and fine-tuning of one's methods. Below is a list of known limitations of the pilot experiment and, wherever possible, some solutions to overcome them in the main experiment:

- Small number of subjects. This is a common problem in translation process research, and we hope the inclusion of more subjects (10) in the future will make data statistically more relevant.
- Experience increases over time. The subjects' experience (thus their speed and quality) in working in both environments (B and V) can increase over time, at least as far as post-editing MT is concerned. So the data we are gathering might be representative of performance at the beginning of a learning curve. One solution would be to train translators for some period and measure their performance after some time.
- Few segments. The text chosen as source text had long segments, which obliged us to use only a few segments per type of suggestion. Since we do not want to increase the total word volume of the source texts, we will probably need to choose another article or even another text type.
- Irregular segments. The shortest segment had six words, while the longest had 44, which makes them hardly comparable, as MT is known to work better with segments containing a 'single idea' and worse with long sentences (cf. Thicke 2011). Same solution as above.
- Terminology. The distribution of terms in the source text should be reconsidered for the main experiment. Even though the time used for terminology search was discounted, the time spent within the translation tool was higher when terms were

more complicated. This was partly compensated for by the fact that the type of suggestion for each segment was defined randomly, but in order to eliminate extraneous variations, we will try to remove problematic terms or provide a glossary for them.

- Segment identification. Sometimes it was difficult to identify which segment translators were focusing on, especially in the self-revising phase. Eye tracking is an additional data-collection method that is being considered to help solve this issue.

Additionally, the provenance information was incomplete for Subject1P (no highlighted differences were shown on the screen), due to a temporary bug in the translation software used in the experiment. A solution to this issue was found and the problem was fixed for the experiment with Subject2P.

Conclusion

We set out to investigate whether provenance information about translation suggestions in translation environments that integrate TM and MT has an impact on *speed*, *amount of editing* and *quality*. We ran a pilot experiment with two subjects that translated two 500-word texts in two different environments. Through screen recording and keystroke logging, we measured the time and the number of characters typed for each of five different types of translation suggestions. The final translated texts were assessed for quality by human reviewers. Retrospective interviews completed the data gathering methodology with an aim at obtaining general impressions from the subject translators.

Our data show that the overall speed was not significantly different in the two scenarios, the amount of editing was higher (between 20 and 49 percent) in the 'blind' environment, and the quality was of comparable level. If we look into individual types of suggestions, data on speed also show that translators spent much longer translating (post-editing) exact matches when they did not know the provenance of the suggestions.

Although inconclusive, the results of the current study indicate that provenance information is relevant for translators working with translation suggestions from TM and MT, and that this information should be taken into account when analysing and comparing the results of different experiments.

Expected benefits

The results of this pilot experiment have given some hints on translation and post-editing processes. The answer we are ultimately trying to find is which strategy is more effective: either the pre-translation of files, with further review at the end, or a 'live' translation with knowledge of the provenance of suggestions. A deeper understanding of both processes, as envisaged in my forthcoming doctoral research, can be beneficial for all parties involved in the translation scene, including independent translators, translation agencies, translation-tool developers and, ultimately, translation customers, as the results can contribute to devise optimal workflows and best practices.

Besides the potential impact on earnings (and savings), the search for optimal processes can increase the volume of text that can be processed. We must remember that in the European Union, as well as in many international organizations, large amounts of text remain untranslated due to time or budget constraints. Even more important, it is our concern to try to optimise the translation process in ways that will help increase job satisfaction among translation professionals. Finally, I hope the results will also be of intellectual importance, as we are trying to demonstrate that the impact of technology is not just in what it does, but also in what the stakeholders know about what it does.

Further questions for future research

A common concern that arises from studies on translation technology in general and on machine translation in particular is whether the technology will replace human translators. As happened at the transitions from pencil to typewriter, from typewriter to computer, from text editor to translation memory, it is our belief that a potential increase in productivity from the integration of TM and MT will not affect occupation levels or incomes, thanks to an increase in translation volumes.

However, this project would benefit from complementary studies on aspects such as the cognitive, ergonomic, sociological and ethical implications of the use of technology. As mentioned above, some questions need to be addressed: How do the new work methods affect job satisfaction? If the participation of non-trained players in mainstream translation (e.g. crowd-sourcing, fan-subbing, etc.) is facilitated and becomes more commonplace through the use of technology, what impact will this have on the translation professional and on the image of professional translators? These are all questions that need to be answered in parallel with research on translation technology as a means to gain efficiency.

References

Allen, Jeffrey. 2003. "Post-editing." In: Harold Somers (ed.) *Computers and Translation: A Translator's Guide*. Amsterdam & Philadelphia: Benjamins. 297-317.

Allen, Jeffrey. 2005. "What is Post-editing?" *Translation Automation Newsletter* 4: 3-5.

Christensen, Tina Paulsen & Schjoldager, Anne. 2010. "Translation-memory (TM) research: what do we know and how do we know it?" *Hermes – Journal of Language and Communication Studies*: 89-101.

Dimitrova, Birgitta Englund. 2005. *Expertise and Explicitation in the Translation Process*. Amsterdam: Benjamins.

Dragsted, Barbara. 2004. *Segmentation in Translation and Translation Memory Systems: An empirical investigation of cognitive segmentation and effects of integrating a TM system into the translation process*. Doctoral thesis, Copenhagen Business School: Samfundslitteratur.

García, Ignacio. 2010. "Is machine translation ready yet?" *Target* 22 (1): 7-21.

Guerberof, Ana. 2002. *Productivity and quality in the post-editing of outputs from translation memories and machine translation*. Minor dissertation. Tarragona: Universitat Rovira i Virgili.

Guerberof, Ana. 2009. "Productivity and quality in the post-editing of outputs from translation memories and machine translation." *Localisation Focus – The International Journal of Localisation* 7/1: 11-21.

Guerra Martínez, Lorena. 2003. *Human translation versus machine translation and full post-editing of raw machine translation output*. Master's dissertation. Dublin: Dublin City University.

Hutchins, W. John & Somers, Harold L. 1992. *An introduction to machine translation*. London: Academic Press.

Jakobsen, Arnt Lykke. 2002. "Translation drafting by professional translators and by translation students." Empirical Translation Studies: Process and Product. *Copenhagen Studies in Language* 27: 191-204.

Jakobsen, Arnt Lykke. 2003. "Effects of think aloud on translation speed, revision, and segmentation." In: Fábio Alves (ed.) *Triangulating Translation. Perspectives in process oriented research*. 69-95.

Krings, Hans. 2001. *Repairing Texts* [edited by G.S. Koby, translated from German by G.S. Koby, G.M. Shreve, K. Mischerikow and S. Litzer]. Kent, Ohio: Kent State University Press. (1994 post-doctoral thesis)

Künzli, Alexander. 2006. "Translation revision – A study of the performance of ten professional translators revising a technical text." In: Maurizio Gotti & Susan Sarcevic (eds.) *Insights into specialized translation*. Bern/Frankfurt: Peter Lang, 195-214.

Künzli, Alexander. 2007. "Translation Revision: a study of the performance of ten professional translators revising a legal text." In: Y. Gambier, M. Shlesinger & R. Stolze (eds.) *Translation Studies: doubts and directions*. Amsterdam: Benjamins. 115-126.

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

Mossop, Brian. 2007. "Empirical studies of revision: what we know and need to know." *The Journal of Specialised Translation* 8, unpaginated.

O'Brien, Sharon. 2002. "Teaching post-editing: A proposal for course content." Proceedings of the European Association of Machine Translation. Available from http://mt-archive.info/EAMT-2002-OBrien.pdf. Visited June 2011.

O'Brien, Sharon. 2006. "Eye-tracking and translation memory matches." *Perspectives: Studies in Translatology* 14/3: 185-205.

Pavlović, Nataša. 2007. *Directionality in collaborative translation processes – A study of novice translators*. Doctoral thesis. Tarragona: Universitat Rovira i Virgili.

Pym, Anthony. 2011a. "Translation research terms. A tentative glossary for moments of perplexity and dispute." In: Anthony Pym (ed.) *Translation Research Projects 3*. Tarragona: Intercultural Studies Group. 75-110.

Pym, Anthony. 2011b. “Democratizing translation technologies – the role of humanistic research.” Paper presented to the Luspio Translation Automation Conference, Rome, April 5 2011.

Thicke, Lori. 2011. “Improving MT results. A study.” *Multilingual* Jan-Feb 2011. 37-41.

Webb, Lynn E. 1998. *Advantages and Disadvantages of Translation Memory*. Master’s dissertation. Monterey: Monterey Institute of International Studies.

Yamada, Masaru. 2011. “The effect of translation memory databases on productivity.” In: Anthony Pym (ed.) *Translation research projects 3*. Tarragona: Intercultural Studies Group. 63-73.

---

[i] Our environment B presents all the pre-translated text at once, while Guerberof’s environment displays each pre-translated segment at a time and does not allow for a revising phase. In order to make both environments closer (and environment B closer to environment V), we are planning to change the settings of our Trados project in the future to have it display each segment at a time. There are no plans to restrict the revising phase in our study.

[ii] Subject1 received the instructions orally, but then the researcher decided to give similar instructions in written format to Subject2, in order to eliminate potential variations due to his oral performance. In the main experiment, all subjects should receive the same instructions in written format.

Appendix 1: Pre-assessment questionnaire

1a. ¿Tienes experiencia con herramientas de memoria de traducción?
1b. ¿Tienes experiencia con alguna versión de Trados? ¿Cuál la más reciente?

2a. ¿Alguna vez has utilizado sugerencias de traducción automática en tus traducciones?
2b. ¿Por qué?

3. ¿Tienes experiencia con la traducción de material de marketing relacionado con áreas técnicas?

4a. ¿Te consideras un traductor profesional? ¿Por qué?
4b. ¿Desde cuándo trabajas como traductor?
4c. ¿Cuál es tu formación académica?

5a. ¿Traduces del inglés al español (castellano)?
5b. ¿Consideras el español tu idioma nativo (materno o paterno)?
5c. ¿Te consideras hablante nativo de otro(s) idioma(s)?

Appendix 2: Text to be copied from Spanish in Word

Los planes para la mayor central solar del mundo, el Blythe Solar Power Project, han recibido la luz verde de la Comisión de Energía de California. Una vez acabado, el proyecto, situado en el desierto de Mojave, tendrá una capacidad de 1.000 megavatios, equivalente a la producción de una central nuclear o una central térmica de carbón moderna. La autorización definitiva llegó en diciembre y ya se ha empezado el trabajo inicial. Su promotor es Solar Millennium LLC y las dos primeras instalaciones se conectarán a la red en 2013 y 2014. Arnold Schwarzenegger, el anterior gobernador de California, aplaudió la decisión de la comisión de autorizar la construcción del proyecto y afirmó: “Proyectos como éste son el futuro de la economía californiana”.

Appendix 3: Text to be translated from English into Spanish in Word

“When it comes to mass production we are still at the research and development stage,” says Francis Richt, who works with composite development at Sandvik Coromant. “But we are counting on this new material soon being used to reduce the weight of electric and hybrid cars.” Richt adds that the appliances in the aerospace industry are more complex than in vehicles, with a greater necessity for quality and with simultaneous processing of composites and other materials such as titanium.

Appendix 4: Source text 1

Per-Ivar Sellergren, development manager at the Volvo Cars Materials Centre in Sweden, is also optimistic. "Everything has gone according to our plans, and we will have a prototype in the form of a car boot by the end of 2012," he says. Cost is an issue, but Sellergren says that even though it is still considerably more expensive than steel and aluminium, the future still lies in composites for electric and hybrid cars.

According to Volvo's calculations, the cost of a bonnet made of the new battery material could be equal to that of an original bonnet plus a lithium-ion battery. "As manufacturers, we could add on an extra amount for the carbon fibre bonnet, because we are, in effect, getting a battery for nothing," he says.

According to Ulf Carlund, Volvo Cars' composite expert, until now production methods have been too slow, and earlier investments in traditional auto plants have needed to be exploited. Partly, it is also due to the fact that steel automakers have found it difficult to think and work with composites and new materials. However, there is a strong will to change, and the public will see more and more polymer materials on the inside and outside of new cars, according to Volvo experts.

Audi, by virtue of the aluminium car A2, is a forerunner in lightweight car manufacturing. In the company's "lightweight" centre in Neckarsulm, Germany, Audi engineers build on new carbon fibre techniques already used by subsidiary Lamborghini as well as on composite technology/expertise developed by parent company Volkswagen's luxury Bugatti model.

The manufacturing process needs to be improved. In the Audi R8 Spyder sports car, which costs more than 120,000 euros and is only made at a rate of about 15–20 per day, Audi uses carbon fibre-strengthened polymer in both the sides and the top of the roof box. One prerequisite that would make it more cost-effective in cheaper mass-produced cars is that a number of aluminium components could be replaced by a single carbon fibre component. "Instead of five or six different tools, maybe you would only need a single tool," says Karl Durst, a development engineer at Audi's Leichtbauzentrum.

Here, among other projects, fibres are packed into a composite material to increase the weight advantage in comparison with aluminium, from around 17–18 percent to about 25 percent. The project hinges on a material that has the same drag and press weight burden capability as aluminium. Despite this, there are still several major and minor problems to solve, says Durst, not least the corrosion in the joints between composites and other materials. There is also a noise factor. For every kilogram that the car becomes lighter, the noise level increases, requiring insulation, which in turn adds more weight. One of the biggest challenges, however, will be the material's familiarity among car mechanics handling it. "It should be possible to fix the car and replace composite car parts in even the smallest Audi workshop anywhere in the world," says Durst.

Appendix 5: Source text 2

Lars Herbeck, who is manager of German machine manufacturer Voith's subsidiary Voith Composites, foresees a large need in several areas. One is with the optimized flow processes for new materials, and another is a paced production of around 100,000 components a year, as well as a much faster pace in the cycle. Compared with aluminium components, which can be made every second, it can take from 20 minutes to an hour for larger composite parts.

Oliver Geiger, who is a researcher in the composite materials department at research institute Fraunhofer-Institut für Chemische Technologie in Pfinztal, Germany, is looking at ways to get large companies to work together in various sectors. Audi's Durst sees the need for a leap forward in the technology, rather than relying on a slower evolutionary development.

Daimler too, which has used carbon fibre in its racing car, the SLR McLaren, since 2004, is also concentrating heavily on developing the technology. In April 2010, it started a cooperation with Japanese chemicals company Toray, the world's leading manufacturer of carbon fibre. The aim is that within three years the company will be developing components made of carbon fibre for models with an average manufacturing volume of 20,000 to 40,000 cars a year.

Meanwhile archrival BMW is being considerably braver. Together with German partner SGL Carbon, BMW is investing 100 million US dollars in a composite factory in Moses Lake, Washington, in the United States. According to BMW Head of Finance Friedrich Eichiner, the factory will make "large volumes at competitive prices" for the first time. The aim is to reduce the price of materials to less than half of the current price of carbon fibre, which is used today in racing cars at a cost of USD 22 to 55 a kilogram.

The carbon fibre will be made on two lines, with an annual capacity of some 1,500 tonnes, and will be used to make the new BMW electric car, the Megacity Vehicle, a four-seat hatchback. A sports car variation, with a small additional diesel engine and two electric engines should be capable of reaching a top speed of more than 200 kilometres an hour.

Megacity is expected to roll off the production line in 2013–2014 in Leipzig, where BMW will be investing more than EUR 400 million in the next couple of years. According to BMW, it will be the world's first production-line car with an entire passenger cell made of light carbon-fibre composite on an aluminium chassis. The first sketches that BMW has released reveal a car that looks straight out of a science fiction movie. It has a battery like a flat mattress under the whole coupe, over-dimensioned wheels and a dynamic, almost aggressive image.

It remains to be seen what the effect will be on the factory floor of an industry already under pressure. "It's a gamble," says a light-weight specialist at one of BMW's competitors. Nevertheless, manager Norbert Reithofer is fully aware of the risks.

Appendix 6: Translation instructions provided for environment V

<u>Instrucciones</u>

Una editorial o una agencia de traducción tienen un contrato para traducir la revista de una gran empresa. Están buscando a otro traductor para remplazar al traductor actual, que ha anunciado que no podrá seguir traduciendo. Has aceptado hacer una prueba, que consiste en un texto de unas 500 palabras.

La empresa de traducción tiene una memoria de traducción, que ha sido creada a partir de la versión final, aprobada por el cliente, de números anteriores de la misma revista. **La empresa de traducción te envía el archivo original, dicha memoria de traducción y te pide para traducir el archivo en la versión de Trados que te han facilitado, configurada de una manera específica. Tendrás a tu disposición las sugerencias de la memoria de traducción (con "fuzzy matches" entre 70% y 100%) y sugerencias del sistema de traducción automática Google Translate.**

Tu tarea consiste en traducir el texto que te han enviado, pensando que después el resultado de tu trabajo será valorado por un revisor profesional, que te va a dar una nota. Para esta prueba, te van a pagar 0,10 € por palabra, independientemente de los "fuzzy matches".

Límite de tiempo: 1,5 hora.

Appendix 7: Translation instructions provided for environment B

Instrucciones

Una editorial o una agencia de traducción tienen un contrato para traducir la revista de una gran empresa. Están buscando a otro traductor para remplazar al traductor actual, que ha anunciado que no podrá seguir traduciendo. Has aceptado hacer una prueba, que consiste en un texto de unas 500 palabras.

La empresa de traducción tiene una memoria de traducción, que ha sido creada a partir de la versión final, aprobada por el cliente, de números anteriores de la misma revista. **La empresa de traducción quiere probar con una herramienta de traducción automática y te envía el texto original y un texto que ha sido pretraducido al español utilizando dicha memoria de traducción (con fuzzy matches entre 70% y 100%) y Google Translate, sin revisión humana.**

Tu tarea consiste en traducir el texto que te han enviado, pensando que después el resultado de tu trabajo será valorado por un revisor profesional, que te va a dar una nota. Para esta prueba, te van a pagar 0,10 € por palabra.

Límite de tiempo: 1,5 hora.

Appendix 8: Tracking of times and edits for each of the segments translated by Subject1 in Environment V

SEGMENT NUMBER	TYPE OF SUGGESTION	1st RENDITION				2nd RENDITION			
		TIME			KEYS & CLICKS	TIME			KEYS & CLICKS
		Start	End	Total		Start	End	Total	
1	FUZZY 76%	02:25.44 04:00.00 07:24.11	03:06.55 06:03.78 07:45.22	41.11 123.78 21.11	- 56 11	00:34.89	00:42.67	07.78	-
2	FUZZY 93%	07:51.33	08:57.33	66.00	9	00:43.44	01:02.44	19.00	-
3	EXACT	09:00.00 12:26.89	09:20.00 12:29.89	20.00 03.00	- -	01:03.22	01:26.11	22.89	8
4	MT	12:29.89 13:51.44	12:51.89 15:27.22	22.00 95.78	3 47	01:26.78	01:49.33	22.55	-
5	FUZZY 94%	15:28.78	17:05.00	96.22	13	01:49.89	02:01.89	12.00	-
6	MT	17:05.78 19:20.00	18:59.89 21:08.44	114.11 108.44	17 92	02:02.56	02:42.22	39.66	-
7	MT	21:19.89	22:59.22	99.33	49	02:42.89	02:56.00	13.11	3
8	EXACT	23:00.44	23:48.11	47.67	-	02:56.89	03:17.33	20.44	-
9	MT	23:48.11 26:22.00 27:03.56	23:56.89 26:30.78 27:23.22	08.78 08.78 19.66	2 - 12	03:18.11	03:23.44	26.55	-
10	EXACT	27:24.44	27:50.44	26.00	-	03:24.22	03:43.22	19.00	-
11	FUZZY 85%	27:50.44	28:29.11	38.67	61	03:43.89 04:02.33	04:02.33 04:27.33	18.44 25.00	- -
12	FUZZY 97%	28:30.44	29:42.67	72.23	8	04:27.33	05:37.22	69.89	75
13	FUZZY 73%	29:43.78	32:03.22	139.44	55	05:38.11	06:42.22	64.11	-
14	EXACT	32:04.67	32:33.00	28.33	-	06:43.00	07:04.33	21.33	3
15	FUZZY 77%	32:33.00	33:48.67	75.67	38	07:05.00	07:21.22	16.22	-
16	MT	33:49.89 35:02.44	34:39.00 35:40.00	49.11 37.56	- 9	07:21.89	07:41.11	19.22	-

		36:28.44 36:42.78 14.34 37:07.22 37:10.78 03.56 37:30.78 37:34.11 03.33 40:17.22 40:21.89 04.67 41:43.11 41:44.44 01.33 46:55.33 47:27.44 32.11 50:10.00 50:22.22 12.22 50:29.56 51:12.67 43.11	- - - - - 51 37 32		
17	MT	51:13.78 52:09.00 55.22	42	07:41.78 07:47.78 06.00	-
18	EXACT	52:10.11 52:40.44 30.33	-	07:48.44 07:58.33 09.89	-
19	FUZZY 84%	52:40.44 53:23.67 43.23 54:59.44 55:49.33 49.89	- 71	07:59.00 08:04.89 05.89	-
20	MT	55:50.44 56:24.89 34.45 57:27.56 57:42.33 14.77	8 -	08:05.55 08:10.22 04.67	-
21	FUZZY 82%	57:43.33 58:04.89 21.56	13	08:11.00 08:13.44 02.44	-

Appendix 9: Tracking of times and edits for each of the segments translated by Subject1 in Environment B

SEGMENT NUMBER	TYPE OF SUGGESTION	1st RENDITION				2nd RENDITION			
		TIME			KEYS & CLICKS	TIME			KEYS & CLICKS
		Start	End	Total		Start	End	Total	
1	FUZZY 75%	00:00.00	00:40.33	40.33	11	18:38.44	18:43.89	05.45	-
2	FUZZY 86%	00:40.56	01:37.78	57.22	39	18:44.56	18:46.22	01.66	-
3	MT	02:30.33	04:31.67	121.34	17	18:46.89	19:19.33	32.44	19
4	MT	04:31.67 05:05.78 06:28.56 11:51.33 14:04.33	04:38.22 05:09.56 06:51.44 13:06.33 14:35.11	06.55 03.78 22.88 75.00 30.78	- - - 60 34	19:20.00	19:28.22	08.22	-
5	MT	14:35.67	15:57.22	81.55	22	19:28.78	19:43.67	14.89	8
6	MT	15:57.89	17:17.89	80.00	48	19:44.33	19:59.89	15.56	-
7	FUZZY 87%	17:18.56	19:14.44	115.88	101	20:00.44	20:16.56	16.12	5
8	EXACT	19:14.44 22:47.22	20:49.11 23:03.78	94.67 16.56	119 1	20:17.22	20:34.33	17.11	9
9	MT	23:04.33 26:08.44 27:53.56	23:24.44 26:52.00 28:15.22	20.11 43.56 21.66	- 68 10	20:35.00	20:59.67	42.45	26
10	FUZZY 95%	28:16.00 31:03.56 31:51.33 33:23.67	30:35.11 31:39.78 32:48.67 34:28.00	139.11 36.22 57.34 64.33	26 - 20 48	21:00.44	21:11.67	11.23	-
11	FUZZY 99%	34:28.56	34:51.67	23.11	40	21:12.22	21:13.33	01.11	-
12	FUZZY 74%	34:52.33 35:41.11 37:46.44 43:11.56 55:10.67	35:14.56 37:17.89 38:04.33 43:19.56 55:51.89	22.23 96.78 17.89 08.00 41.22	- 90 2 - 38	21:14.11	21:25.11	11.00	-
13	EXACT	55:52.44	57:12.22	79.78	50	21:25.78	21:39.56	13.78	-
14	EXACT	57:12.78	57:23.22	10.44	3	21:40.00	21:44.33	04.33	-

		57:35.89 58:25.22 49.33 59:10.22 59:31.89 21.67	35 -		
15	MT	59:32.56 00:44.89 72.33	36	21:44.89 21:58.78 13.89	-
16	EXACT	00:45.56 01:27.44 41.88 02:28.78 02:48.22 19.44 05:22.44 05:33.56 11.12 05:37.78 06:25.33 47.55 06:54.00 07:17.89 23.89 09:11.78 09:31.00 19.22 11:05.67 11:24.22 18.55 12:07.89 12:28.11 20.22 13:15.67 13:32.11 16.44 13:44.22 14:24.11 39.89	- - - - - - 10 - - 75	21:59.44 22:04.11 04.67	-
17	FUZZY 86%	14:24.67 15:01.44 36.77	21	22:04.67 22:17.11 12.44	-
18	MT	15:02.00 15:14.00 12.00	11	22:17.67 22:18.89 01.22	-
19	FUZZY 93%	15:14.67 15:35.22 20.55 15:56.00 16:24.00 28.00	31 44	22:19.44 23:00.44 41.00	23
20	FUZZY 72%	16:24.56 17:11.56 47.00	39	23:01.11 23:04.33 03.22	-
21	EXACT	17:12.22 17:47.11 34.89	9	23:04.89 23:14.44 09.55	-

Appendix 10: Tracking of times and edits for each of the segments translated by Subject2 in Environment V

SEGMENT NUMBER	TYPE OF SUGGESTION	1st RENDITION				2nd RENDITION			
		TIME			KEYS & CLICKS	TIME			KEYS & CLICKS
		Start	End	Total		Start	End	Total	
1	FUZZY 76%	04:26.00	07:58.92	212.92	135	41:26.61	41:47.77	21.16	6
2	FUZZY 93%	08:00.31 09:36.31	09:16.92 10:31.15	76.61 54.84	24 28	41:48.69	43:02.62	73.93	43
3	EXACT	10:32.46	11:28.31	55.85	13	43:03.38	43:18.38	15.00	
4	MT	11:29.92	13:01.00	91.08	40	43:18.92	43:20.46	01.54	
5	FUZZY 94%	13:02.23	13:53.23	51.00	6	43:21.69	44:02.85	41.16	
6	MT	13:54.46 15:10.15	14:42.00 15:57.69	47.54 47.54	36 13	44:02.85 46:21.85 47:25.38 48:26.46	44:13.31 46:34.46 47:36.15 48:32.08	10.46 12.61 10.77 05.62	12
7	MT	15:59.69	16:47.00	47.31	48	48:32.08	49:03.08	31.00	
8	EXACT	16:51.54	17:22.46	30.92	7	49:04.69	49:12.85	08.16	
9	MT	17:23.23	17:46.77	23.54	2	49:13.61 50:41.54 52:35.38	49:16.08 50:48.54 52:39.69	11.39 07.00 04.31	
10	EXACT	17:49.54	18:30.54	41.00	13	52:42.38	53:06.61	24.23	
11	FUZZY 85%	18:31.15	19:40.92	69.77	53	53:06.92	53:41.54	34.62	7
12	FUZZY 97%	19:42.15	22:33.77	171.62	261	47:48.46 53:41.77	47:58.31 54:16.85	09.85 35.08	13 8
13	FUZZY 73%	22:35.38 25:14.46 26:55.31	23:15.69 25:53.77 27:48.46	40.31 39.31 53.15	 5 10	54:18.08	55:00.77	42.69	
14	EXACT	27:54.77	28:44.31	49.54	3	55:01.08	55:29.85	28.77	
15	FUZZY 77%	28:44.85 30:20.00 31:32.38	29:55.85 30:35.69 31:56.00	71.00 15.69 23.62	75	55:33.54 01:21.15	56:33.38 01:32.92	59.84 11.77	18 11

16	MT	32:11.69 34:20.38 128.69	166	56:33.62 56:49.00 15.38 57:14.61 57:49.31 34.70	34
17	MT	34:24.00 34:57.00 33.00	12	57:49.54 58:02.23 12.69	
18	EXACT	34:59.23 35:40.92 41.69 36:19.08 36:35.85 16.77		58:02.46 58:14.23 11.77 58:32.85 58:53.54 20.69	
19	FUZZY 84%	36:36.38 37:22.00 45.62 38:22.23 39:10.31 48.08	14 26	58:53.77 59:02.31 08.54	
20	MT	39:11.46 40:07.46 56.00	49	59:02.54 59:59.00 56.46	26
21	FUZZY 82%	40:08.77 41:10.77 62.00		00:01.38 00:35.62 34.24	16

Appendix 11: Tracking of times and edits for each of the segments translated by Subject2 in Environment B

SEGMENT NUMBER	TYPE OF SUGGESTION	1st RENDITION				2nd RENDITION			
		TIME			KEYS & CLICKS	TIME			KEYS & CLICKS
		Start	End	Total		Start	End	Total	
1	FUZZY 75%	08:14.54 09:07.15	08:49.92 09:12.92	35.38 05.77	13 -	[?]	[?]	13.385	-
2	FUZZY 86%	09:13.15	10:28.69	75.54	85	[?]	[?]	13.385	-
3	MT	10:28.92 12:55.31	11:28.08 13:00.77	59.16 05.46	- 5	05:18.92	05:32.69	13.77	5
4	MT	11:31.38 12:49.38 13:00.77 14:28.85 14:49.00 15:59.46	11:56.38 12:54.15 14:25.46 14:37.31 15:02.85 16:08.23	25.00 04.77 84.69 08.46 13.85 08.77	- - 54 - 15 -	05:34.31	05:36.77	02.46	-
5	MT	16:08.46	17:15.08	66.62	62	05:40.85	06:00.00	19.15	2
6	MT	17:15.31	18:34.08	78.77	54	06:00.00	06:23.54	23.54	-
7	FUZZY 87%	18:34.31	20:24.92	110.61	153	06:25.85	08:23.31	117.46	155
8	EXACT	20:25.15 22:20.61	22:06.38 22:28.54	101.23 07.93	109 -	08:23.31	08:33.15	09.84	
9	MT	22:28.77	23:07.15	38.38	32	08:37.46	09:08.00	44.69	20
10	FUZZY 95%	23:07.46 24:10.15	23:27.46 26:13.31	20.00 123.16	- 76	09:08.00	09:43.23	35.23	-
11	FUZZY 99%	26:13.54	26:32.69	19.15	13	09:43.23	09:50.69	07.46	-
12	FUZZY 74%	26:35.62 29:26.00 30:25.15 35:42.15	28:30.23 29:33.62 30:59.23 35:56.69	114.61 07.62 34.08 14.54	112 - 6 10	09:52.23	10:28.38	36.15	11
13	EXACT	35:56.92	38:26.38	149.46	139	10:28.38	10:52.46	24.08	-
14	EXACT	38:32.23 42:39.46	39:04.38 42:48.23	32.15 08.77	- -	10:53.62	11:05.00	11.38	-

15	MT	42:51.92 45:34.31 46:21.38 50:43.92	43:38.15 45:48.54 46:41.92 51:56.54	46.23 14.23 20.54 72.62	10 - - 57	11:07.92	11:53.38	45.46	21
16	EXACT	51:56.85 52:32.62 55:46.77 56:28.69 58:54.54	52:07.31 53:36.31 55:58.46 56:42.08 59:12.31	10.46 63.69 11.69 13.39 17.77	- 13 - 17 -	11:57.08	12:27.23	30.15	15
17	FUZZY 86%	59:17.92	59:57.31	39.39	37	12:27.69	12:39.85	12.16	-
18	MT	59:57.54	00:18.15	20.61	36	12:40.77	12:53.54	12.77	6
19	FUZZY 93%	00:18.31 02:11.23 02:48.15	01:40.54 02:28.77 03:00.62	82.23 17.54 12.47	88 - 19	12:53.54	13:03.15	09.61	3
20	FUZZY 72%	03:00.85	04:17.38	76.53	82	13:04.00	13:33.23	29.23	20
21	EXACT	04:17.61	04:46.09	28.48	13	13:33.23	13:52.54	19.31	2

Appendix 12: Speed and amount of editing consolidated per type of suggestion for Subject1 in Environment V

		SOURCE WORDS	1st RENDITION TIME (sec)	1st RENDITION SPEED (words/h)	2nd REND. TIME (sec)	COMBINED SPEED (words/h)	TARGET CHARS	1st RENDITION KEYS & CLICKS	1st RENDITION AMOUNT OF EDITING	2nd REND. KEYS & CLICKS	COMBINED AMOUNT OF EDITING
EXACT (100%) MATCHES	**SEGMENT #1**	23	23	3600	22.89	1804	164	0	0.0%	8	4.9%
	SEGMENT #2	33	47.67	2492	20.44	1744	171	0	0.0%	0	0.0%
	SEGMENT #3	25	26	3462	19	2000	162	0	0.0%	0	0.0%
	SEGMENT #4	29	28.33	3685	21.33	2102	154	0	0.0%	3	1.9%
	SEGMENT #5	21	30.33	2493	9.89	1880	141	0	0.0%	0	0.0%
	SUB-TOTAL	**131**	**155.33**	**3036**	**93.55**	**1895**	**792**	**0**	**0.0%**	**11**	**1.4%**
90-99% MATCHES	**SEGMENT #1**	33	66	1800	19	1398	190	9	4.7%	0	4.7%
	SEGMENT #2	21	96.22	786	12	699	104	13	12.5%	0	12.5%
	SEGMENT #3	37	72.23	1844	69.89	937	170	8	4.7%	75	48.8%
	SUB-TOTAL	**91**	**234.45**	**1397**	**100.89**	**977**	**464**	**30**	**6.5%**	**75**	**22.6%**
80-89% MATCHES	**SEGMENT #1**	21	38.67	1955	18.44	1324	159	61	38.4%	0	38.4%
	SEGMENT #2	20	93.12	773	5.89	727	123	71	57.7%	0	57.7%
	SEGMENT #3	10	21.56	1670	2.44	1500	94	13	13.8%	0	13.8%
	SUB-TOTAL	**51**	**153.35**	**1197**	**26.77**	**1019**	**376**	**145**	**38.6%**	**0**	**38.6%**
70-79% MATCHES	**SEGMENT #1**	21	186	406	7.78	390	114	67	58.8%	0	58.8%
	SEGMENT #2	36	139.44	929	64.11	637	195	55	28.2%	0	28.2%
	SEGMENT #3	30	75.67	1427	16.22	1175	148	38	25.7%	0	25.7%
	SUB-TOTAL	**87**	**401.11**	**781**	**88.11**	**640**	**457**	**160**	**35.0%**	**0**	**35.0%**
NO MATCHES (MT FEEDS)	**SEGMENT #1**	37	117.78	1131	22.55	949	265	50	18.9%	0	18.9%
	SEGMENT #2	24	222.55	388	39.66	330	167	109	65.3%	0	65.3%
	SEGMENT #3	20	99.33	725	13.11	640	142	49	34.5%	3	36.6%
	SEGMENT #4	7	37.22	677	26.55	395	62	14	22.6%	0	22.6%
	SEGMENT #5	28	201.34	501	19.22	457	173	129	74.6%	0	74.6%
	SEGMENT #6	19	55.22	1239	6	1117	121	42	34.7%	0	34.7%
	SEGMENT #7	15	49.22	1097	4.67	1002	88	8	9.1%	0	9.1%
	SUB-TOTAL	**150**	**782.66**	**690**	**131.76**	**591**	**1018**	**401**	**39.4%**	**3**	**39.7%**
	TOTAL	**510**	**1726.9**	**1063**	**441.08**	**847**	**3107**	**736**	**23.7%**	**89**	**26.6%**

Appendix 13: Speed and amount of editing consolidated per type of suggestion for Subject1 in Environment B

			1st RENDITION		2nd REND.	COMBINED		1st RENDITION		2nd REND.	COMBINED
		SOURCE WORDS	TIME (sec)	SPEED (words/h)	TIME (sec)	SPEED (words/h)	TARGET CHARS	KEYS & CLICKS	AMOUNT OF EDITING	KEYS & CLICKS	AMOUNT OF EDITING
EXACT (100%) MATCHES	SEGMENT #1	30	111.23	971	17.11	842	152	120	78.9%	9	84.9%
	SEGMENT #2	30	79.78	1354	13.78	1154	178	50	28.1%	0	28.1%
	SEGMENT #3	25	81.44	1105	4.33	1049	150	38	25.3%	0	25.3%
	SEGMENT #4	18	258.2	251	4.67	247	156	85	54.5%	0	54.5%
	SEGMENT #5	25	34.89	2580	9.55	2025	156	9	5.8%	0	5.8%
	SUB-TOTAL	**128**	**565.54**	**815**	**49.44**	**749**	**792**	**302**	**38.1%**	**9**	**39.3%**
90-99% MATCHES	SEGMENT #1	38	297	461	11.23	444	308	94	30.5%	0	30.5%
	SEGMENT #2	7	23.11	1090	1.11	1040	41	40	97.6%	0	97.6%
	SEGMENT #3	20	48.55	1483	41	804	119	75	63.0%	23	82.4%
	SUB-TOTAL	**65**	**368.66**	**635**	**53.34**	**555**	**468**	**209**	**44.7%**	**23**	**49.6%**
80-89% MATCHES	SEGMENT #1	27	57.22	1699	1.66	1651	108	39	36.1%	0	36.1%
	SEGMENT #2	24	115.88	746	16.12	655	160	101	63.1%	5	66.3%
	SEGMENT #3	26	36.77	2546	12.44	1902	148	21	14.2%	0	14.2%
	SUB-TOTAL	**77**	**209.87**	**1321**	**30.22**	**1155**	**416**	**161**	**38.7%**	**5**	**39.9%**
70-79% MATCHES	SEGMENT #1	16	40.33	1428	5.45	1258	108	11	10.2%	0	10.2%
	SEGMENT #2	44	186.12	851	11	804	216	130	60.2%	0	60.2%
	SEGMENT #3	17	47	1302	3.22	1219	94	39	41.5%	0	41.5%
	SUB-TOTAL	**77**	**273.45**	**1014**	**19.67**	**946**	**418**	**180**	**43.1%**	**0**	**43.1%**
NO MATCHES (MT FEEDS)	SEGMENT #1	31	121.34	920	32.44	726	219	17	7.8%	19	16.4%
	SEGMENT #2	30	138.99	777	8.22	734	162	94	58.0%	0	58.0%
	SEGMENT #3	26	81.55	1148	14.89	971	153	22	14.4%	8	19.6%
	SEGMENT #4	28	80	1260	15.56	1055	223	48	21.5%	0	21.5%
	SEGMENT #5	15	85.33	633	42.45	423	95	78	82.1%	26	109.5%
	SEGMENT #6	29	72.33	1443	13.89	1211	184	36	19.6%	0	19.6%
	SEGMENT #7	6	12	1800	1.22	1634	33	11	33.3%	0	33.3%
	SUB-TOTAL	**165**	**591.54**	**1004**	**128.67**	**825**	**1069**	**306**	**28.6%**	**53**	**33.6%**
	TOTAL	**512**	**2009.06**	**917**	**281.34**	**805**	**3163**	**1158**	**36.6%**	**90**	**39.5%**

Appendix 14: Speed and amount of editing consolidated per type of suggestion for Subject2 in Environment V

		SOURCE WORDS	1st RENDITION TIME (sec)	1st RENDITION SPEED (words/h)	2nd REND. TIME (sec)	COMBINED SPEED (words/h)	TARGET CHARS	1st RENDITION KEYS & CLICKS	1st RENDITION AMOUNT OF EDITING	2nd REND. KEYS & CLICKS	COMBINED AMOUNT OF EDITING
EXACT (100%) MATCHES	SEGMENT #1	23	55.85	1483	15	1169	165	13	7.9%	0	7.9%
	SEGMENT #2	33	30.92	3842	8.16	3040	177	7	4.0%	0	4.0%
	SEGMENT #3	25	41	2195	24.23	1380	162	13	8.0%	0	8.0%
	SEGMENT #4	29	49.54	2107	28.77	1333	154	3	1.9%	0	1.9%
	SEGMENT #5	21	58.46	1293	44.53	734	141	0	0.0%	0	0.0%
	SUB-TOTAL	**131**	**235.77**	**2000**	**120.69**	**1323**	**799**	**36**	**4.5%**	**0**	**4.5%**
90-99% MATCHES	SEGMENT #1	33	131.45	904	73.93	578	168	52	31.0%	43	56.5%
	SEGMENT #2	21	51	1482	41.16	820	94	6	6.4%	0	6.4%
	SEGMENT #3	37	171.62	776	44.93	615	230	261	113.5%	21	122.6%
	SUB-TOTAL	**91**	**354.07**	**925**	**160.02**	**637**	**492**	**319**	**64.8%**	**64**	**77.8%**
80-89% MATCHES	SEGMENT #1	21	69.77	1084	34.62	724	133	53	39.8%	7	45.1%
	SEGMENT #2	20	93.7	768	8.54	704	104	40	38.5%	0	38.5%
	SEGMENT #3	10	62	581	34.24	374	84	0	0.0%	16	19.0%
	SUB-TOTAL	**51**	**225.47**	**814**	**77.4**	**606**	**321**	**93**	**29.0%**	**23**	**36.1%**
70-79% MATCHES	SEGMENT #1	21	212.92	355	21.16	323	136	135	99.3%	6	103.7%
	SEGMENT #2	36	132.77	976	42.69	739	192	15	7.8%	0	7.8%
	SEGMENT #3	30	110.31	979	71.61	594	151	75	49.7%	29	68.9%
	SUB-TOTAL	**87**	**456**	**687**	**135.46**	**530**	**479**	**225**	**47.0%**	**35**	**54.3%**
NO MATCHES (MT FEEDS)	SEGMENT #1	37	91.08	1462	1.54	1438	261	40	15.3%	0	15.3%
	SEGMENT #2	24	95.08	909	39.46	642	161	49	30.4%	12	37.9%
	SEGMENT #3	20	47.31	1522	31	919	131	48	36.6%	0	36.6%
	SEGMENT #4	7	23.54	1071	22.7	545	60	2	3.3%	0	3.3%
	SEGMENT #5	28	128.69	783	50.08	564	172	166	96.5%	34	116.3%
	SEGMENT #6	19	33	2073	12.69	1497	111	12	10.8%	0	10.8%
	SEGMENT #7	15	56	964	56.46	480	88	49	55.7%	26	85.2%
	SUB-TOTAL	**150**	**474.7**	**1138**	**213.93**	**784**	**984**	**366**	**37.2%**	**72**	**44.5%**
	TOTAL	**510**	**1746.01**	**1052**	**707.5**	**748**	**3075**	**1039**	**33.8%**	**194**	**40.1%**

Appendix 15: Speed and amount of editing consolidated per type of suggestion for Subject2 in Environment B

		SOURCE WORDS	1st RENDITION TIME (sec)	1st RENDITION SPEED (words/h)	2nd REND. TIME (sec)	COMBINED SPEED (words/h)	TARGET CHARS	1st RENDITION KEYS & CLICKS	1st RENDITION AMOUNT OF EDITING	2nd REND. KEYS & CLICKS	COMBINED AMOUNT OF EDITING
EXACT (100%) MATCHES	SEGMENT #1	30	109.16	989	9.84	908	182	109	59.9%	0	59.9%
	SEGMENT #2	30	149.46	723	24.08	622	182	139	76.4%	0	76.4%
	SEGMENT #3	25	40.92	2199	11.38	1721	155	0	0.0%	0	0.0%
	SEGMENT #4	18	117	554	30.15	440	143	30	21.0%	15	31.5%
	SEGMENT #5	25	28.48	3160	19.31	1883	156	13	8.3%	2	9.6%
	SUB-TOTAL	**128**	**445.02**	**1035**	**94.76**	**854**	**818**	**291**	**35.6%**	**17**	**37.7%**
90-99% MATCHES	SEGMENT #1	38	143.16	956	35.23	767	319	76	23.8%	0	23.8%
	SEGMENT #2	7	19.15	1316	7.46	947	42	13	31.0%	0	31.0%
	SEGMENT #3	20	112.24	641	9.61	591	122	107	87.7%	3	90.2%
	SUB-TOTAL	**65**	**274.55**	**852**	**52.3**	**716**	**483**	**196**	**40.6%**	**3**	**41.2%**
80-89% MATCHES	SEGMENT #1	27	75.54	1287	13.39	1093	105	85	81.0%	0	81.0%
	SEGMENT #2	24	110.61	781	117.46	379	184	153	83.2%	155	167.4%
	SEGMENT #3	26	39.39	2376	12.02	1821	152	37	24.3%	0	24.3%
	SUB-TOTAL	**77**	**225.54**	**1229**	**142.87**	**752**	**441**	**275**	**62.4%**	**155**	**97.5%**
70-79% MATCHES	SEGMENT #1	16	41.15	1400	13.39	1056	97	13	13.4%	0	13.4%
	SEGMENT #2	44	170.85	927	36.15	765	235	128	54.5%	11	59.1%
	SEGMENT #3	17	76.53	800	29.23	579	125	82	65.6%	20	81.6%
	SUB-TOTAL	**77**	**288.53**	**961**	**78.77**	**755**	**457**	**223**	**48.8%**	**31**	**55.6%**
NO MATCHES (MT FEEDS)	SEGMENT #1	31	64.62	1727	13.77	1424	224	5	2.2%	5	4.5%
	SEGMENT #2	30	145.54	742	2.46	730	158	69	43.7%	0	43.7%
	SEGMENT #3	26	66.62	1405	19.15	1091	143	62	43.4%	2	44.8%
	SEGMENT #4	28	78.77	1280	23.54	985	222	54	24.3%	0	24.3%
	SEGMENT #5	15	38.38	1407	44.69	650	96	32	33.3%	20	54.2%
	SEGMENT #6	29	153.62	680	45.46	524	198	67	33.8%	21	44.4%
	SEGMENT #7	6	20.61	1048	12.16	659	30	36	120.0%	6	140.0%
	SUB-TOTAL	**165**	**568.16**	**1045**	**161.23**	**814**	**1071**	**325**	**30.3%**	**54**	**35.4%**
	TOTAL	**512**	**1801.8**	**1023**	**529.93**	**790**	**3270**	**1310**	**40.1%**	**260**	**48.0%**