Long Distance Revisions in Drafting and Post-editing

Carl, Michael; Kay, Martin; Jensen, Kristian Tangsgaard Hvelplund

Document Version
Final published version

Publication date:
2010

Long Distance Revisions in Drafting and Post-editing

Michael Carl[1,2], Martin Kay[1] Kristian T.H. Jensen[2]

[1] Stanford University, H-STAR and Department of Linguistics
[2] Copenhagen Business School, Languages & Computational Linguistics, Frederiksberg, Denmark

Abstract. This paper investigates properties of translation processes, as observed in the translation behaviour of student and professional translators. The translation process can be divided into a *gisting*, *drafting* and *post-editing* phase. We find that student translators have longer gisting phases whereas professional translators have longer post-editing phases. Long-distance revisions, which would typically be expected during post-editing, occur to the same extent during drafting as during post-editing. Further, both groups of translators seem to face the same translation problems. We suggest how those findings might be taken into account in the design of computer assisted translation tools.

1 Introduction

In contrast to the large number of publications on MT post-editing, little research has been carried out on how translators review and post-edit their own translations. Lörscher[10], one of the pioneers in translation process research, points out:

> Solving translation problems is often carried out as a series of steps. Generally, subjects do not immediately reach solutions which they consider to be optimal. ... subjects generally use (linguistically) simple strategies first, and only when they turn out to be unsuccessful do the subjects employ more complex strategies. This procedure of the subjects complies with the generative principle whereby complex translation strategies are ... derived from simpler structures. (p:430)

Revision and post-editing of drafted translation are thus in order and indicative of the complexity (or uncertainty) of a translation problem. Only few years ago, research on human translation processing was based on think-aloud protocols [4,9,10], however, recent technological developments have made it possible to directly analyse user activity data (UAD), notably eye movement data and keystroke data [5,3].

In a recent study, Malkiel [11] investigates the predicatability of "self-revisions" in English-Hebrew translations, based on manual analysis of

the revision keystrokes. In this paper, we use our triangulation technology [2,3] and discuss a method to automatically detect and analyse revision patterns.

Given the increasing interest in interactive Machine Translation, [8,13][3] and in the design of man-machine interfaces, we expect that insights derived from the study of human translation processing will provide valuable information for the designers of MT post-editing tools.

2 Gisting, Drafting and Post-editing

We base our research on a translation experiment [7] in which 12 professional and 12 student translators produced translations using the Translog [5] software.[4] Translog presents the source text (ST) in the upper part of the monitor, and the target text (TT) is typed in a window in the lower part of the monitor. When the start button is pressed, the ST is displayed and eye movement and keystroke data are registered. The task of the translator is then to type the translation in the lower window. After having completed the translation, the subject presses a stop button, and the translation, along with the translation process data, are stored in a log file.

Translators vary greatly with respect to how they produce translations. However, the process can be divided into three phases, which we refer to as *gisting*, in which the translator acquires a preliminary notion of the ST, *drafting* in which the actual translation is typed (drafted), and *post-editing* in which some or all of the drafted text is re-read, typos corrected and sentences rearranged or reformulated on the background of the translator's better understanding of the text by the time this stage is reached.

2.1 Translation Progression Graphs

The UAD can be represented in so-called translation progression graphs[12]. Figure 1 shows translation progression graphs for two students (S17 and S23) at the top and at the bottom respectively and a professional (P1) in the middle. The graphs plot activity data which was collected during the translation of a 160 words text from English into Danish.[5]

[3] Google has just made available a toolkit for human assisted translation with more than 50 languages.

[4] The software can be downloaded from `www.translog.dk`

[5] The English source text is shown in the Appendix.

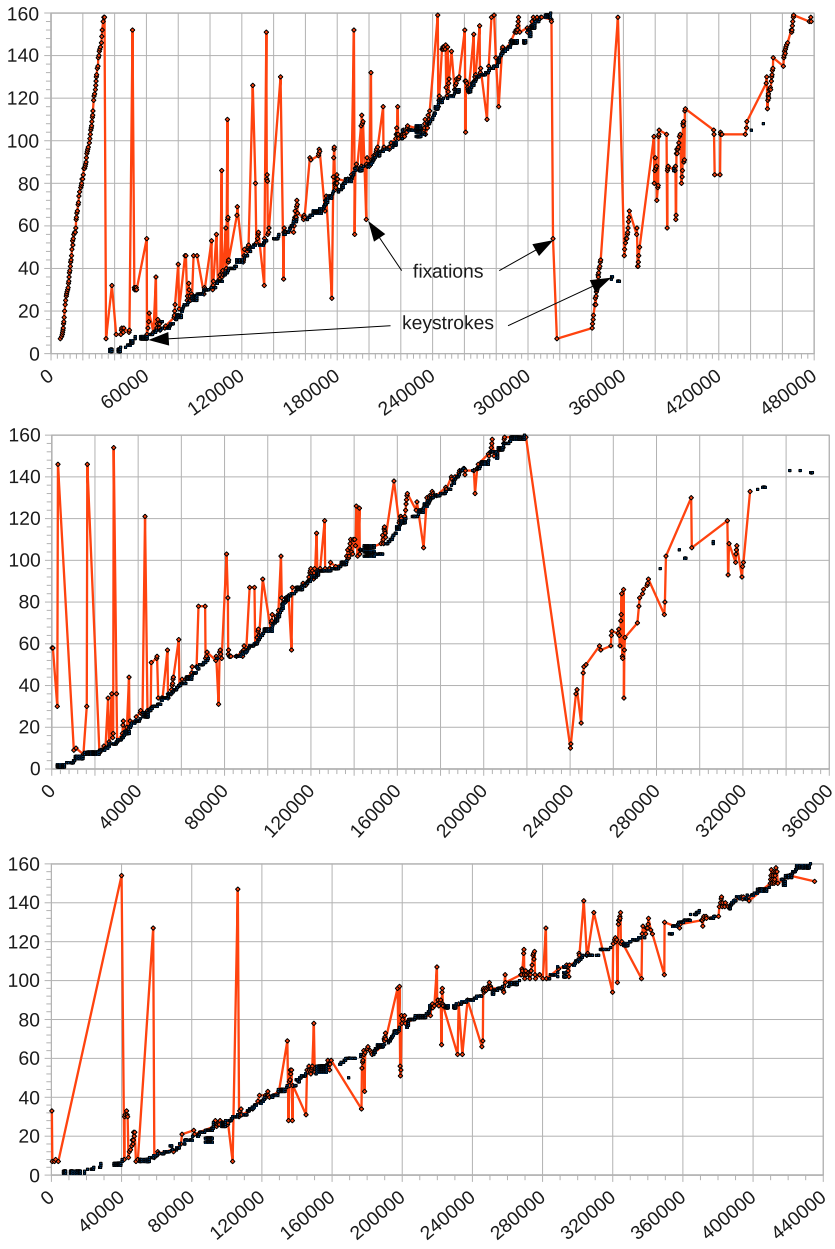

Fig. 1. Three translation progression graphs from top down subjects S17, P1 and S23, showing keystrokes and eye movements: S17 shows a clear distinction into gisting, drafting and post-editing. P1 has no gisting phase and spends almost 50% of the translation time on post-editing, while S23 only has a drafting phase.

The horizontal axis represents the translation time in milliseconds, and the vertical axis represents the source-language words from the beginning of the text (bottom), to the end (top). As described in Carl, 2009 [2], keystrokes that contribute to the TT, are mapped onto the ST words which they translate. All keystrokes that contribute to the translation of the ith source word are represented as single dots in the ith line from the bottom of the graph. The red (i.e. grey) line plots the gaze activities on the source text words. Single eye fixations are marked with a dot on the fixation line[6].

The progression graph of subject S17 (top graph in figure 1) shows a clear distinction between gisting, drafting and post-editing. Subject S17 spends almost 40 seconds getting acquainted with the text. The graph shows the progression of fixations nicely in which the ST is apparently read from beginning to end.

The drafting phase takes place between seconds 40 and 320. Eye movements can be observed where the translator moves back and forth between the ST and the TT. Some fixations are captured during this journey between the current ST position and the TT window (or to the keyboard) which are mapped on text positions remote from the current location of the corresponding translation.

The drafting phase is followed by a post-editing phase, from approx. second 320 until second 480. Translator S17 seems to re-read much of the ST during post-editing, but only few keystrokes occur, i.e. around seconds 360 and 440.

Translator P1, the second graph in figure 1, shows virtually no gisting phase. The first keystrokes can be observed less than 5 seconds after the ST appears on the screen. P1 also has a long post-editing phase of two minutes, from seconds 220 to 360. A number of revision keystrokes are visible, around seconds 300 and 340.

A third translation pattern for translator S23 is shown in final graph. No gisting and no post-editing take place, but some revision occurs at various places, e.g. around seconds 100 and 170. The time it takes to produce the translations is between 6 minutes (P1) and 8 minutes (S17).

2.2 Translation Expertise and Translation Phases

For students, there is a clear tendency towards longer gisting and shorter post-editing phases, whereas professional translators have shorter gisting

[6] Notice that only fixations on the source text are represented in the graph. Our software was not able to compute and map fixations on the emerging target text words.

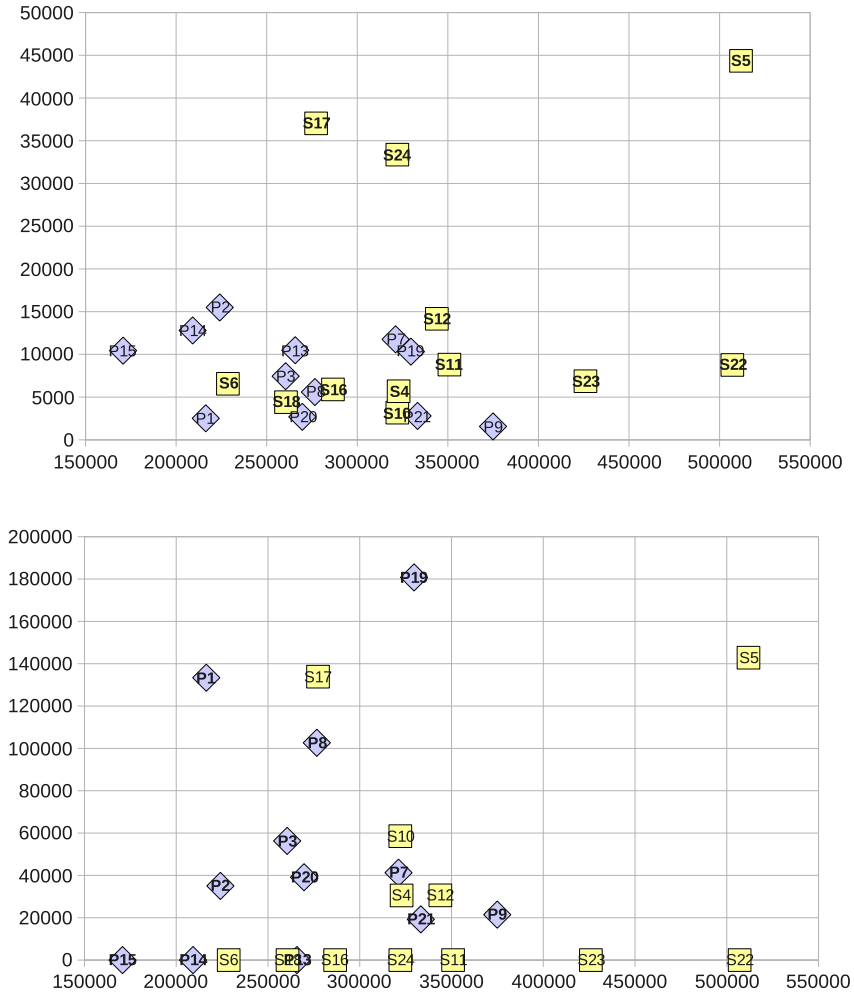

Fig. 2. Top: drafting time (horizontal) and gisting time (vertical). Rectangular symbols represent student translators, diamond shapes represent professionals. Students spend more time on gisting than professionals. Bottom: drafting time (horizontal) and post-editing time (vertical). Rectangular symbols represent students, diamond shapes represent professionals. On average, professionals spend more time post-editing than do students; many students completely skip post-editing.

and longer post-editing phases. Figure 2.1 (top) plots the relationship between drafting and gisting time, and the bottom graph in figure 2.1 shows the relation between drafting time and post-editing time. Almost all professional translators (9 out of 12) engage in some kind of post-editing, while 7 out of 12 student translators do not post-edit. The inverse observation can be made with respect to gisting: 3 students but no professional translator engage in gisting for more than 20 seconds. These results are only partially in line with Jakobsen, 2002 [6] who finds that professional translators invest more time than students in gisting and post-editing, but are faster at drafting the translation.

3 Long Distance Revisions

Changes in the target text translation may take place at any moment during drafting or post-editing: in the middle or at the end of a word or after or at the end of a sentence or paragraph. We distinguish between two types of revisions, short-distance revisions, and long-distance revisions.

3.1 Translation Phases and Long Distance Revisions

Long-distance revisions occur if two successive keystrokes are located 2 or more words apart from each other. For instance, a translator might first translate "nurse" into "sygeplejerske", but when she realizes that the 'nurse' is in fact masculin, she might correct all occurrences into "sygeplejer"[7]. To do so, the cursor must move to a previous words, and if the corrected word is two or more words apart from the last cursor action we will observe long-distance keystrokes. A long-distance revision is thus a sequence of two successive keystrokes, which are located in a different part of the target text translation. All other modifications of drafted text are short-distance revisions. Whereas short-distance revisions most likely are associated with typing errors, which the translator immediately corrects, it is plausible that long distance revisions are indicative of 'real' translation problems that the translator is struggling with.

One would expect that long-distance revisions are particularly abundant during post-editing; however, our data indicate that they occur with the same frequency although no separate post-editing phase takes place.[8] Figure 3 suggests that the post-editing time and the number of long-distance revisions are basically independent: long-distance revisions take

[7] In our classification below this would correspond to a IDWX pattern.
[8] An example of this is subject S23 in figure 1, above.

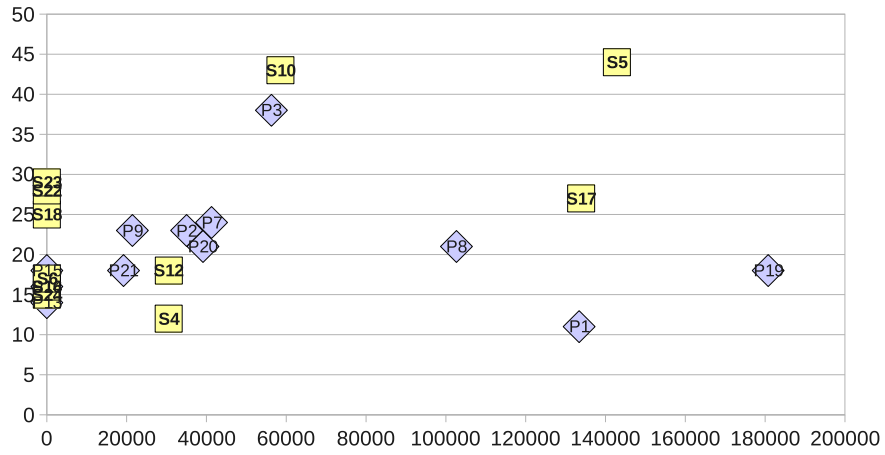

Fig. 3. Number of long distance revisions (vertical) and post-editing time (horizontal) shows the parameters to be unrelated. Long distance revisions occur equally frequently for students as for professionals, irrespectively of the length of the post-editing phase.

place in approximately equal number, whether or not there is a separate post-editing phase. Thus, more experienced, professional translators seem to prefer a modular mode of working, in whicn both types of editing are separated in two clearly different phases. Conversely, students are more likely to mix those two phases. Jakobsen [6] reports similar findings in his experiments, where students produce more revisions during drafting.

Figure 3 also shows that translators perform between 11 and 45 long distance revisions on the 160 word text. Students perform slightly more revisions, on average one revision every 6.5 words, while professionals revise once every 7.8 word. This figure approximately coincides with the one given in Malkiel [11] whose student translators "self-revise" every 8th word. In the next sections we will show that these revision are by no means equally distributed in the text.

3.2 Patterns in Long Distance Revisions

A related question is whether and to what extend translators face the same difficulties during translation. That is, we may be confident that translators share similar problems if long distance revisions cluster at particular text positions so that common patterns can be observed in the UAD.

Indeed, figure 4 shows that revisions of the 24 translators occur more frequently at certain positions in the texts. The graph shows four or five positions where many revision take place, i.e. around word positions 14, 50, 105, 120 and 151. The contexts of these passages are shown in `bold` in the Appendix. We briefly discuss some of the difficulties that these particular passages might present to a translatior.

A word-for-word translation of "imprisoned for life today" would not be idiomatic in Danish. In order to find an idiomatically acceptable rendering of the expression, the translator would have to reorder the constituents and make different lexical choices.

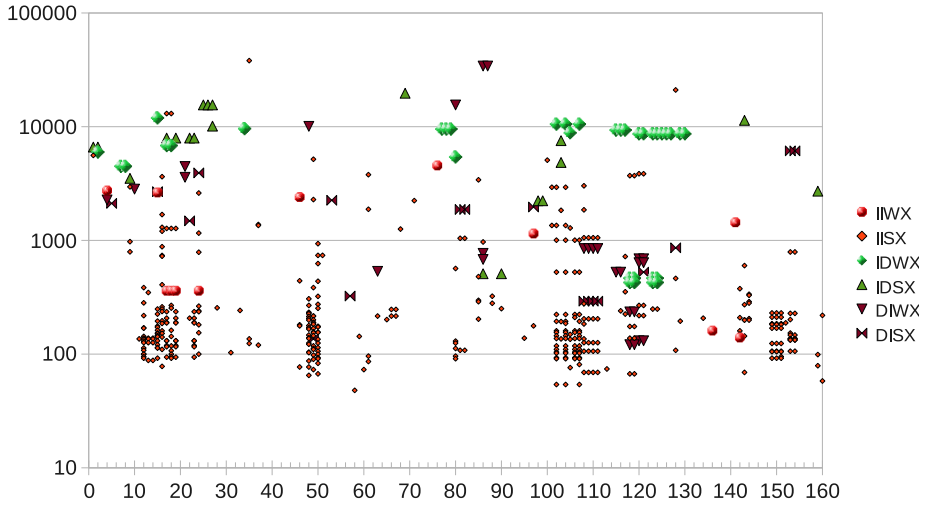

Fig. 4. Elapsed time (vertical) and positions of long distance revisions in the translation (horizontal): The horizontal axis enumerates the source-language words (0 to 160)and the dots in the graph represent different types of long distance revisions of their translations.

The translation of "counts of murder" into Danish may cause difficulty since the expression occurs infrequently in this context. The translator would have to test several Danish equivalent expressions in order find an acceptable one. The translation data shows more than 12 possible solutions for this passage.

The compound expression "hospital staff" has no exact equivalent in Danish. The translator would have to test several possible translation

alternatives before reaching a satisfying solution. This difficulty can also be measured by the fact that the data contain 20 different translations for "awareness of other hospital staff".

3.3 Classifying Long Distance Revisions

The keystrokes in our representations can be either text-inserting or text-deleting. That is, keystrokes for mere cursor movement are skipped and ignored in the graphs. Accordingly, in order to classify the long-distance revisions, we distinguish between insertion (I) and deletion (D) revisions. Since each of the two keystrokes in a revision can be an insertion or a deletion, we have four categories of pairs of revision keystrokes. In addition, we also distinguish the situation in which the second keystroke immediately follows a word separator (S) from the situation in which the second keystroke is in the middle of a word (W). Thus, in principle there could be eight types of long-distance revision.[9] The six most frequent combinations are shown in figure 4 and are briefly described below: [10]

1. IISX: two successive long-distance insertion keystrokes, the second immediately following a word separator, e.g. inserting an article.
2. IIWX: two successive long-distance insertion keystrokes, the second not immediately following a word separator, e.g inserting a suffix of a word.
3. IDSX: an insertion followed by a long-distance deletion keystroke occurring at the beginning of a word, e.g. deleting an article.
4. IDWX: an insertion followed by a long-distance deletion keystroke occurring in the middle of a word, e.g deleting a suffix.
5. DISX: a deletion followed by a long-distance insertion that occurs at the beginning of a dislocated word, e.g. inserting an article.
6. DIWX: a deletion followed by a long-distance insertion in the middle of a dislocated word, e.g inserting a suffix of a word.

Table 1 summarizes revision types for all 24 translations. It gives rise to the following observations: revisions usually start at the beginning of a word (461 occurrences) and less frequently in the middle (74 occurrences). ID revision patterns require much more time than DI or II revisions. That is, the time lapse between the end of an insertion and the beginning of a

[9] The long-distance between successive keystrokes is marked as X in the examples below.

[10] Unfortunately, our data show too few instances for 'DD' revisions to draw any conclusions.

deletion in another passage of the text is much higher than that between a deletion to a following insertion, or two successive insertions. On average, the pause between the insertion and the long-distance deletion is 7734ms and 8676ms respectively for the deletion to take place at the beginning and the middle of a word while it is only a fraction of this for the other types of revisions.

Type	IIWX	IISX	IDWX	IDSX	DIWX	DISX
Number of occurrences	12	423	35	20	27	18
Average time interval	755	169	8676	7734	689	1677

Table 1. Number of occurrences and time interval between the two keystrokes of for several types of long-distance revision pattern.

Presumably, the reason for the long ID revision is that a meaning hypothesis was realized and finished by the last insertion, and a new meaning hypothesis must mature before the deletion can take place. This would require much more anticipation and effort than a DI pattern, where the long-distance insertion is presumably only a consequence of the thought that lead to the deletion, or for the II patterns where the second insertion is a continuation of the first insterion.

4 Conclusion

Three phases can be distinguished in human translation: a *gisting phase*, a *drafting phase* and a *post-editing phase*. In our relatively short and simple text, gisting and post-editing seem to be optional: professional translators skip the gisting phase, tend to start immediately with drafting and have a longer post-editing phase. Novices, in contrast require a longer gisting phase, and often completely skip post-editing. In line with this, [7] finds that students "allocate considerably more time to each ST segment", our investigation indicates that this might be due to the longer gisting phase.

However, there seems to be an equal number of long-distance revisions for students and professionals. Hence, students revise parts of their translations when drafting, while professional translators work more structured and postpone revisions to a post-editing phase. Interestingly, irrespectively of when the revision is made, students and professionals revise the same parts of the translations, presumably because they face the same problems in the translation.

In order to figure out which of the phases in a translation process can be mechanized, computer assistance might be conceived to support the translator's structuring of the following task: gisting support tools could prepare the translator for difficulties of the ST, giving them e.g. a review of frequently used terms in their contexts, point to unusual collocations, etc., whereas translation memories or MT post-editing tools [8,13] might be a basis for drafting and post-editing support.

Special attention in the design of automated support during drafting and post-editing should receive the ID revision patterns, where translators spend much of their time here.

If certain translation and post-editing strategies turn out to be more successful than others, as in the case of our professional translators, then they should presumably be taken into account in the design of translation support tools. Under this assumption, a MT post-editing tool seems to be better grounded than a translation completion tool [1], which would mix drafting and post-editing phases, as we have observed in novice translators.

References

1. Sergio Barrachina, Oliver Bender, Francisco Casacuberta, Jorge Elsa, Civera Cubel, Shahram Khadivi, Antonio Lagarda, Hermann Ney, Jesús Tomás, Enrique Vidal, and Juan-Miguel Vilar. Statistical Approaches to Computer-Assisted Translation. *Computational Linguistics*, pages 3–28, 2009.
2. Michael Carl. Triangulating product and process data: quantifying alignment units with keystroke data. *Copenhagen Studies in Language*, 38:225–247, 2009.
3. Michael Carl and Arnt Lykke Jakobsen. Towards statistical modelling of translators' activity data. *International Journal of Speech Technology*, 12(4):124–146, 2010.
4. P. Gerloff. Second Language Learners Reports on the Interpretive Process: Talk-aloud Protocols of Translation. In *[?]*, pages 243–262, 1986.
5. Arnt Lykke Jakobsen. Logging target text production with Translog. In *[?]*, pages 9–20, 1999.
6. Arnt Lykke Jakobsen. Translation drafting by professional translators and by translation students. In *[?]*, pages 191–204, 2002.
7. Kristian T. H. Jensen. Distribution of attention between source text and target text during translation. In *IATIS*, 2009.
8. Philipp Koehn and Barry Haddow. Interactive Assistance to Human Translators using Statistical Machine Translation Methods. In *MT Summit*, 2009. `http://www.mt-archive.info/MTS-2009-TOC.htm`.
9. H. Krings. *Was in den Köpfen von Übersetzern vorgeht*. Gunter Narr, Tübingen, 1986.
10. W. Lörscher. Investigating the Translation Process . *Meta*, XXXVII(3):426–439, 1992.

11. Brenda Malkiel. From Ántona to My Ántona: tracking self-corrections with Translog. volume 37 of *Copenhagen Studies in Language*, pages 149–167. Copenhagen: Samfundslitteratur, 2008.
12. Daniel Perrin. Progression analysis (PA): investigating writing strategies at the workplace. *Pragmatics*, 35:907–921, 2003.
13. Marco Trombetti. Creating the World's Largest Translation Memory . In *MT Summit*, 2009. http://www.mt-archive.info/MTS-2009-TOC.htm.

Appendix: Source Test

Killer nurse receives four life sentences
Hospital Nurse Colin Norris was **imprisoned for life today for the killing of four of his patients**. 32 year old Norris from Glasgow killed the four women in 2002 by giving them large amounts of sleeping medicine. Yesterday, he was found **guilty of four counts of** murder following a long trial. He was given four life sentences, one for each of the killings. He will have to serve at least 30 years. Police officer Chris Gregg said that Norris had been acting strangely around the hospital. Only the **awareness of other hospital staff put a stop to him** and to the killings. The police **have learned that the motive** for the killings was that Norris disliked working with old people. All of his victims were old weak women with heart problems. **All of them could be considered** a burden to hospital staff.