

# Ape: Reducing the Monkey Business in Post-Editing by Automating the Task Intelligently

Claus Povlsen and Annelise Bech\*

Center for Sprogteknologi  
Njalsgade 80, DK-2300 Copenhagen S  
Denmark  
[claus@cst.ku.dk](mailto:claus@cst.ku.dk)

\* Lingtech A/S  
Vesterbrogade 24  
DK-1620 Copenhagen V  
[abe@lingtech.dk](mailto:abe@lingtech.dk)

## Abstract

For a professional user of MT, quality, performance and cost efficiency are critical. It is therefore surprising that only little attention – both in theory and in practice - has been given to the task of post-editing machine translated texts. This paper will focus on this important user aspect and demonstrate that substantial savings in time and effort can be achieved by implementing intelligent automatic tools. Our point of departure is the PaTrans MT-system, developed by CST and used by the Danish translation company Lingtech. An intelligent post-editing facility, Ape, has been developed and added to the system. We will outline and discuss this mechanism and its positive effects on the output. The underlying idea of the intelligent post-editing facility is to exploit the lexical and grammatical knowledge already present in the MT-system's linguistic components. Conceptually, our approach is general, although its implementation remains system specific. Surveys of post-editor satisfaction and cost-efficiency improvements, as well as a quantitative, benchmark-based evaluation of the effect of Ape demonstrate the success of the approach and encourage further development.

## Keywords

Machine translation, user aspects, post-editing.

## Introduction

Machine translation is now a serious alternative to manual translation. Many organisations and businesses employ MT-systems and for various purposes. Some use MT-systems for information purposes (gisting of material in “exotic” languages), some as a basis for decision-making (as to which documents to provide high-quality translations of) and some use machine translation in their production (providing an output to be post-edited and finalised before “publication”/delivery to the client).

Lingtech, a professional translation company in Copenhagen, Denmark is one of the pioneers in using MT in its production of high-quality technical translations. Since 1993, Lingtech has used PaTrans<sup>1</sup> to translate technical texts, primarily patents, and currently some 3.5 million words are run through the system every year.

The output from the MT-system is post-edited and finalised by Lingtech staff before a final translation is delivered to the client. Thus, Lingtech has a considerable interest in the time and money spent on this task.

(Previously, Lingtech has reported on improved cost-benefit from providing ancillary tools and setting up a suitable work-flow for preparing texts for machine translation and extending dictionaries (Bech, 1997)).

From the user point of view, steps to improve and reduce the post-editing workload is not only an issue of profit maximisation, although it is obviously an incentive, which cannot be neglected in the business context. The quality of the final product as well as the ergonomics and job-satisfaction of the post-editors are also at stake.

Over the years, Lingtech has carefully and systematically monitored and registered problem-areas in relation to the so-called peripheral tasks of pre- and post-editing. In close collaboration with the development team at Center for

---

<sup>1</sup> PaTrans was developed by Center for Sprogteknologi for Lingtech. The system translates from English to Danish. The system is described in (Ørsnes et al., 1996; and Maegaard & Hansen, 1995).

Sprogteknologi, Lingtech has striven to continually automate and facilitate the tasks involved.

An interesting recent development is the introduction of an intelligent strategy for minimising the burden of post-editing. A specific programme component, Ape, has been added to the system, which exploits the linguistic knowledge of PaTrans and the grammatical information in the translated text to correct a number of “mistakes”, which would otherwise have had to be dealt with in the manual post-editing.

In the next sections, we will set the context and outline the fundamentals of the intelligent post-editor.

## PaTrans and Post-Editing

PaTrans is a fully automatic, transfer-based machine translation system. The user prepares texts for translation, codes necessary dictionary entries and then submits the text for translation. There is no interaction with the system during the translation process; the user takes over again when the system has finished and produced a translated text. The user’s task is then to post-edit the output in order to produce a flawless final version.

Based on the Eurotra linguistic model, PaTrans has an analysis component (source grammar and lexicon), a transfer component, and a synthesis component (target grammar and lexicon). The linguistic strategy involves processing the input text such that surface neutral representations of the input sentences are produced in terms of word order and function words. The linguistic representation is an ordered predicate argument structure with valency-bound lexical items and function words encoded as information on the relevant nodes. The representation(s) of sentences are transferred to the synthesis component, which produces a surface sentence in the target language.

In order to operate in practice, various robustness features have been implemented in PaTrans to ensure that the system always produces a translation of an input sentence. Input sentences that for various reasons cannot be processed correctly or completely<sup>2</sup> by the linguistic components proper are treated by the failsoft component of the system.

Although the PaTrans system produces output of a quite high quality, the user is unavoidably - as with any MT-system - faced with output that needs to be post-edited.

From the commercial point of view, the time and effort required for getting from translated output to a final version of the translation is crucial. The less post-editing is required or the easier it is to correct flaws, the better the cost-efficiency of the whole operation.

In the case of PaTrans, the output consists of both well-formed translations of sentences according to the linguistic components of the system and of failsofted sentences. For the post-editor, the failsofted output typically requires the greatest attention and effort. These parts of the text often present the post-editor with hard-to-understand sentences or muddled-up word order; that is the most tedious and time-consuming (and hence costly) aspect of post-editing. Not surprisingly, in a controlled survey of post-editing problems conducted by the Lingtech staff, word order problems were ranked as the number one irritation factor in post-editing.

With this in mind, we consequently set out to remedy the situation, the result of which is the intelligent post-editing component Ape.

## Ape: Basic Idea and Functionality

The quality in terms of correct word order in failsofted sentences depends to a large degree on where in the PaTrans translation process the failsoft mechanism has been activated. The lowest translation quality is achieved if failsofting takes place after the predicate argument structure representation of the input sentence has been generated.

An index is assigned to words in the input sentence indicating the order in which they appear and this information is preserved throughout the translation process. After the linguistic processing has finished Ape traverses through the translation results generated and in cases where they have been tagged as failsofted, the translation is re-ordered according to the indices of the content words in the source input sentence.

One could easily think of language pairs in which the idea of using the original word order as a knowledge source for re-ordering the translation results would be inappropriate. The relatively high degree of similarity between English and Danish led to the assumption that implementation of re-ordering would result in significantly improved translation quality. Various differences between English and Danish word order can, however, be observed. The positioning of adverbs is one example:

'Unfortunately, **the secondary chamber is injection moulded** in two halves which has caused several problems'

->

'Uheldigvis **sprøjtestøbes det sekundære kammer** i to halvdele, hvilket har forårsaget flere problemer.'

The example illustrates the inversion phenomenon in Danish. Whenever i.a. an adverb is topicalised, the word order (unlike English) changes so that the finite verb in (main clauses) precedes the subject of the sentence. Even though it would be possible to treat this difference between English and Danish, it was considered to be a minor problem and left out in the first version of the Ape program.

---

<sup>2</sup> E.g. the input may be ungrammatical, lexical items not coded or fall outside of the linguistic coverage of the system.

The overall re-ordering algorithm in the current version of Ape is thus the following:

If the sentence is failsofted  
and  
if the indexing is out of order  
then re-order all the words (except adverbs) in the  
sentence so that it corresponds to the word order in the  
source input sentence.

In addition to information about indexing and failsoftedness, Ape has access to information about part of speech of the words in the translated sentences and their morphology.

The following examples illustrate the Ape functionality, (the words re-ordered are in bold):

The technique using an inert gas to form the secondary chamber and then only forming the orifice as the secondary chamber is placed **into the container**.

->

Without Ape

Teknikken **ind i containeren** der anvender en inert gas for at danne det sekundære kammer og derefter kun at danne munden idet det sekundære kammer placeres

->

With Ape

Teknikken der anvender en inert gas for at danne det sekundære kammer og derefter kun at danne munden idet det sekundære kammer placeres **ind i containeren**.

The post-edited translation

Teknikken, der anvender en inert gas for at danne det sekundære kammer og derefter kun danner munden som det sekundære kammer, placeres inde i containeren.

The copper compounds **used** as anti-oxidants in this invention may be chosen from those **described** in the document as suitable for lubricants.

->

Without Ape

De **anvendte** kobberforbindelser som antioxidanter i denne opfindelse kan vælges fra de i dokumentet som egnet til smøremidler **beskrevet**.

->

With Ape

De kobberforbindelser **anvendte** som antioxidanter i denne opfindelse kan vælges fra de **beskrevet** i dokumentet som egnet til smøremidler.

The post-edited translation

Kobberforbindelserne som er anvendt som antioxidanter i denne opfindelse kan vælges fra dem som er beskrevet i dokumentet som egnede til smøremidler.

As can be seen, Ape has re-ordered the word order so it corresponds to the source input sentence. The word re-ordering via Ape has improved profoundly the translation

quality and at the same time reduced the manual post-editing work<sup>3</sup>.

In particular the re-ordering of *described/beskrevet* via Ape in the second example has made it possible to grasp the overall meaning of the sentence.

## Treatment of indexless items

A transfer-based approach with stepwise refinement has as one of its implications that function words are elevated (featurised) during analysis of the source language (for a thorough description see (EUROTRA 1991)). This refined representation is then transferred to the target language and the necessary function words are then inserted during generation of the target language (synthesis).

Consequently, index information about the position of the source function words is not available and the target function words do not have indices. In the following a brief description of how indexless words are treated in Ape will be given.

The handling of indexless items is done in two ways.

- a) Attachment of specific constructions as a single indexed item before running Ape.
- b) Defining additional constraints in the overall re-ordering algorithm.

The attachment strategy is quite simple. Indexless words are attached to the (possibly distant) following indexed word. One exception to this strategy is the auxiliary verbs, which are attached to the next indexed verb.

The additional constraints on the overall re-ordering algorithm are the following:

If (the current word's index is higher than the following word and the difference between the current word's index and that of the following word is one)  
and  
if (the following word has initial indexless items)  
then output those indexless items before outputting the current word.

As an illustration consider the following example of the extended re-ordering procedure (lower case letters represent indexless words, the numbers are indexed words and underscore expresses attachment):

The original input sentence:

'... adjusting a mechanism ...'

The translation result before running Ape:

'... en(a)\_mekanisme(2) for(b)\_at(c)\_indstille(1) ...'

In this example all the conditions are met. The index of the current word (represented by 2) is higher than the following word (represented by 1) and the difference

<sup>3</sup> Please bear in mind that both sentences are failsofted so the translation results are not expected to be perfect.

between the indices of the current word and of the following word is one, and the following word (1) has indexless items (b,c). The result after running Ape is:

'... for at indstille en mekanisme ...'

The treatment of indexless items is per definition language-specific so the Ape handling of function words is exclusively restricted to Danish. The overall concept of re-ordering failsofted sentences based on the word order of the source input sentence, however, has (cf. below) a more general perspective.

### Evaluation

Two types of evaluation of the effect of Ape have been carried out, reflecting the two convergent, yet complementary, interests of end-user and system developer.

The user-oriented evaluation is predominantly subjective and focuses on the qualitative aspects in terms of how the post-editor experiences the output when Ape has been applied. That is, evaluation indicates the effect on post-editor satisfaction and task ergonomics.

Complementing this user-oriented evaluation, is the calculation of the effect on cost-efficiency – an important parameter in the commercial setting.

Thus, Lingtech conducted focused interviews with a number of post-editors. The conclusion was – not surprisingly – that the post-editors felt that the number of “difficult” sentences had reduced considerably.

Lingtech also performed a calculation of improvements in cost-efficiency by way of comparing the average number of post-edited words before and after the introduction of Ape. Again, the results were encouraging, in that performance improvements between 10-15% were recorded. In other words, after the introduction of Ape, the overall time spent on getting from machine translated output to finalised text was reduced.

In order to generate a quantitatively more precise measure of the Ape functionality, the following evaluation procedure is being performed. First, a representative corpus of failsofted sentences was identified and collected. Then parallel translation results were made by running PaTrans with and without Ape, respectively<sup>4</sup>. Both these sets of results are then compared automatically with the post-edited and thus final versions of the translation, which function as a benchmark representing the satisfaction rate. The comparison procedure goes through the parallel translation results and aligns them with the benchmark results and then a comparison of the performance of the two PaTrans systems is made. In this way an approximate measure of the Ape functionality in

terms of re-ordering adequacy and thus reduced post-editing is achieved.

### Conclusion

When using machine translation in the process of producing translations of publishable quality, post-editing (i.e. correcting flaws and mistakes in the machine translated output) is an important task that needs to be seriously focused on. As has been argued previously in Bech 1997 and also in this paper, cleverly automating peripheral tasks and providing suitable tools to support the human tasks to be performed in relation to the commercial usage of machine translation are critical parameters for success.

In this paper we have presented an innovative approach to easing the burden of post-editing, going beyond providing an environment with pre-implemented short-cut key operations for the repetitive types of corrections to be made by the post-editor. The basic idea of the Ape strategy is to exploit the linguistic information present in the text to ‘repair’ different kinds of flaws which are tedious to deal with manually in post-editing. We have demonstrated the viability of the strategy by its practical implementation in PaTrans and its positive effect on output quality as reviewed by post-editors (qualitative evaluation) and in terms of a benchmark-based, quantitative evaluation.

As any serious, modern MT-system exploits linguistic knowledge in its processing, the fundamental idea behind the approach presented here is generalisable to other MT-systems and scenarios. Finally, the encouraging results we have obtained with our present version of Ape lead us to work on future developments and further enhancements.

### References

- Bech, A. (1997). MT from an Everyday User’s Point of View. In Proceedings of MT Summit VI (pp. 98--105). San Diego, CA: AMTA.
- EUROTRA (1991). Copeland, C., Durand, J., Krauwer, S. & Maegaard, B. (Eds.), Studies in Machine Translation and Natural Language Processing, Vol. 1. Luxembourg: CEC.
- Maegaard, B. & V. Hansen (1995). PaTrans – Machine Translation of Patent Texts. From Research to Practical Application. In Convention Digest: Second Language Engineering Conference (pp. 1--8). London.
- Povlsen, C., Underwood N., Music B., Neville A. (1998). Evaluating Text-type Suitability for Machine Translation a Case Study on an English-Danish MT System. In Proceedings of the First International Conference on Language Resources & Evaluation (pp. 27--32). Granada.
- Ørsnes, B., B. Music and B. Maegaard (1996). PaTrans – A Patent Translation System. In Proceedings of COLING, (pp. 1115--1118). Copenhagen.

<sup>4</sup> In this evaluation setup we think of PaTrans as being two systems, PaTrans with and PaTrans without the Ape functionality.



