

ALPHA



White paper

---

# Fair Pricing of Machine Translation Projects

## Alpha

St Andrew's House  
St Andrew's Road  
Cambridge CB4 1DL  
United Kingdom

[@thisisalphalive](https://twitter.com/thisisalphalive)

[thisisalpha.com](https://thisisalpha.com)

Alpha processes millions of machine-translated words every year. Some projects involve post-editing files that have been machine-translated by the Client, while other projects are machine-translated at Alpha, using our own installation of Moses. Irrespective of the origin of the project, the big question remaining is whether the MT output is useful, and if so, how useful?

This is a subjective matter, and two post-editors working on the same project and language may have differing opinions. It is therefore important to convert a subjective notion into a quantifiable answer. Many evaluation methods exist (BLEU, METEOR, Edit Distance), but they produce a score that is difficult to interpret.

Such scores may be useful when determining whether the quality of the MT output is improving or deteriorating, but it is virtually impossible to translate them into something as tangible as a word rate. Indeed, these are the two hardest questions to answer:

- What is an appropriate rate to charge for Machine Translation?
- How should we assess whether the rate applied by the Client corresponds to the effort made by the post-editors?

---

## A new evaluation method

Alpha has developed its own evaluation method, known as Reverse Analysis. The Reverse Analysis is carried out using the same CAT tools as used by the post-editors and is based on comparing standard analysis logs, and extrapolating a weighting.

This weighting is the percentage of the no-match rate that should be applied to the MT output. It is not based on time, but on words. We deliberately made this choice in order to counter the fact that post-editing MT output is an emotional task (those who do not enjoy it will probably work more slowly than those who do), but also one that involves different cerebral gymnastics. A study conducted by Dublin City University showed that being experienced in translation did not necessarily result in a high post-editing throughput. Indeed, some students with no translation background performed better than seasoned translators. Simply put, some people are more suited to post-editing than others.

*The aim of the Reverse Analysis process is:*

- To determine the percentage of the MT output that has been edited, and to what extent.
- To determine how to charge for MT content (calculating the weighting that should be applied to the MT match category).

*The Reverse Analysis process involves three steps:*

1. Analysing the source text against an empty TM to determine the word count, the number of chargeable (effective) words, and the Source Compression Ratio.
2. Analysing the machine-translated text against a reversed TM which contains the final, post-edited segments as the source language and the English text as the target language.

*The word count is used to calculate the number of chargeable (effective) words and the Target Compression Ratio. The word count also tells us to what degree the final output matches the MT output, i.e. how much was left unedited (meaning the MT output was perfect), how much was edited a little (meaning the MT output qualifies as a high match, 85-99%) and how much output was edited to such an extent that it bears little or no resemblance to the original MT output (meaning the MT output qualifies as a below 74% or No match).*

### **3. Comparing the Source and Target Ratios and determining the weighting which should be applied to the MT output.**

*This weighting tells us how to charge for MT words. If the output was very close to the final version, a lower charge should naturally be applied; whereas a very different output has required considerable post-editing and should be charged at a higher rate.*

---

## **Considerations**

Because the quality of MT output may vary from project to project (even for the same Client and language), it is important to carry out Reverse Analyses at regular intervals and assess the quality of MT output over time. This means that Alpha can have constructive discussions with Clients, whether they provide machine-translated files for post-editing or rely on Alpha to produce the MT output.

---

Reverse Analysis does not highlight the recurring issues encountered by the post-editors. It is important to gather linguistic feedback and to optimise the MT engine in order to improve the quality of MT output for future projects.

---

### **Found that interesting?**

If you're not already signed up to our newsletter, email [marketing@thisisalpha.com](mailto:marketing@thisisalpha.com) to receive lots more great articles that will help you on your journey to going global.