Machine Translation Acceptance
Among Professional Linguists:
Are We Nearing the Tipping Point?

Author   yves@champollion.net
                                                                     ceo, Wordfast LLC, Paris, France

Abstract

This paper evaluates the level of acceptance of Machine Translation (MT) among translators as of Summer, 2016. Translators are notoriously shy about MT, and their level of acceptance has a lot to say about the level reached by MT.

1. Credits

This text format is derived from the AMTA guidelines for documents submitted for inclusion in AMTA proceedings.

2. MT Among Translators: the Early Days

Translators have been aware of Machine Translation (MT) ever since it first began: the era of automated weather bulletin translation by robots after World War II. But they remained outsiders for a few decades. The Nineties can be seen as the first foray of MT in the world of translators, when the first off-the-shelf solutions (Systran, PowerTranslator, ProMT, etc.) appeared at affordable prices. Still, computer savviness was a prerequisite, which greatly limited the spread of MT use among translators.

3. The Real Start

The nineties witnessed a few earth-shaking changes that redesigned the life of translators for ever, like the widespread acceptance and use of wordprocessors. The spread of the internet is another key factor. While the Internet (and online sales) are credited for the downfall of middlemen like travel agencies, the same factor can be credited for the opposite in translation: the sharp increase in the number of translation agencies. With the ability to team up hundreds of translators and quickly form ad hoc translation task forces for large translation projects, the translation agency, as we know it today, was born.

One collateral effect of agencies attracting the bulk of translation orders was the relentless push toward lower costs. Human translation was considered a relative luxury before the nineties. With the age of luxury for the masses (from smoked salmon to ocean cruises to oversea vacations) came translation for the masses, at least in the business world. Businesses and companies, however small they are, can today call an agency and get 100 pages translated into 20 languages in a week: that was a surhuman feat in the past, available only to very large organizations.

4. Was MT Here to Stay?

The nineties taught us that one low-costing any human activity is how addictive that particular activity becomes. Spend one vacation abroad on an exotic island and presto, you must get back there every year. The addiction of the masses to luxury inexorably leads to a price war, where every trick is authorized. The translation industry is not immune to that effect. Much to the contrary: it is a textbook example of that trend.

As of 2016, we see the maturity of the translation-for-the-masses model. Little more can be added to the existing array of tricks and techniques that agencies use to satisfy a translation-addicted business world.

However, just as air transport or ocean cruises, the translation industry has reached a pricing floor below which it cannot plunge.

Or can it?

5. MT's Age of Reason

Machine Translation is one factor that can prolong the downward spiral of prices in the translation industry. With a twist.

In other industries, when a price war is declared and low-cost is the order of the day, quality is sacrificed first. Summer fruits and vegetables are now available at dirt-cheap prices even in Winter - but they taste like wet cardboard. Senior members of the audience remember a glorious and distant past when tomatoes where had only in season, but Lord did they taste good.

On the contrary, MT's salvation lies in the increase of its quality - because unlike tomatoes, quantity is no issue at all.

As hinted earlier, MT is one of the factors, perhaps the only one, that the translation industry can use to prolong the price war, to keep lowering costs. The human factor cannot be compressed any further - the price paid to translators cannot be significantly lowered at this stage.

The acceptance of MT by linguists lies at the heart of the equation.

6. MT Acceptance Among Translators: Where we Come From

In the 2000's, ten years ago, MT usage was still between low and non-existent among translators, remaining at the level it was in the nineties. A comprehensive research made at London's Imperial College (Elina Lagoudakis, 2006[1]) surveyed Translation Memory usage (MT) among translators, in short, technology used in the translation process, as TM was the prevalent technology back then. The expression "Machine Translation" only appears *once* in 36 pages - and actually does not refer to MT as we discuss it. Those were the days! Machine... what?

That does not mean MT was not used before 2010. But it was rare among independent freelance translators. However, translation agencies would sometimes inject MT into pre-processed translation jobs, but that was not a widespread practice, and remained undercover. Linguists raised eyebrows at MT usage, and many would outright refuse; there was still as strong stigma associated with artificial translation. Another survey on translation technology for individual translators (Luciano Monteiro, 2009[2]) was published by

---

[1] Elina Lagoudakis, in "*Translation Memories Survey*" published by the Imperial College of London, 2006.
[2] Luciano Monteiro, in "*Translators Now And Then - How Technology Has Changed Their Trade*", www.proz.com, 2009.

www.proz.com. The article has a belated entry for Machine Translation (stuck just behind "Other useful software"), which mentions that "*when coupled with terminology management, and post-editing services, MT can provide an attractive cost/benefit solution*". MT was still, as late as 2009, seen by translators as a minor, last-resort crutch for those who needed speed.

7. MT Acceptance Among Translators: Today

Hard statistics about translators' habits are hard to come by. My focus here is on translators defined as *individual* practitioners: freelance translators, and employed translators who have a say in their workflow. The reason statistics are hard to collect is that translators are very scattered. The profession is atomized into indidivual, isolated, practitioners.

To make things more difficult, agencies are shy about revealing their real practices, the technology they use, their prices. Prying reliable information out of translators and agencies is not easy, and will certainly be obsolete in a short few years.

One category of translation tools is the online Computer-Assisted Translation (CAT) tool: a browser-based alternative to the classical, installed CAT tool that translators love to hate. The online CAT tool is on the upswing, especially among two classes of translators: the younger generation, and translators in emergent markets. Whence a precaution about the following figures: the surveyed population is not characteristic of the entire population of translators, as of 2016. But biology and economics being what they are, that young and emergent population will inevitably become mainstream.

8. Today in Figures

Statistics are a difficult to handle properly, and can mean just about anything. Stats on the acceptance of MT by translators are difficult to form. We can only formally poll the use of MT among translators; as for acceptance, which is an attitude toward MT, we can only get clues.

The stats below are derived from two formal sources and one informal source. The two formal sources are an online translation tool (Wordfast Anywhere) with a community of 25,000 registered users, and over 3,000 regular users translating for over ten hours every month. In that situation, figures are reliable, as the tool provides detailed stats on the setup, as well as MT consumption, for each connected translator. The other source is derived from an installed tool (Wordfast Classic and Wordfast PRO), and the associated hotline, which registers the nature of hotline calls, and therefore has a good overview on MT usage.
The last source, an informal one, is the speaker's personal experience as a former translator and project manager, a trainer, a CEO in the translation industry, and a CAT evangelist. While not incorporated in the figures, that experience was used to perform sanity checks on the figures, and to offer an interpretation of the figures.

A. MT usage from paid sources
Paid sources are basically subscription-based MT providers, the ubiquitous ones being Microsoft Translator and Google Translate, but there are others, like iTranslate4.eu. We should note that most paid sources cost literally nothing per month for a typical freelancer's consumption: about the price of a good beer. Still, the need to fill a form and provide credit card details ensures that users are 1. indeed professional translators, and 2. deliberately opt for MT.

The stats here is: 15% of translators use a paid source. The statistics in Wordfast Anywhere use IP numbers to track the approximate location of translators, and it appears that most of those using paid MT are in Europe (45% of the grand total), followed by North America (30%). The rest is evenly distributed around the world.
Stats in installed tools use email addresses, language code, and hotline call records to estimate location, and they concur with the above figures.

B. MT usage from free sources

Free MT sources are numerous; we may quote WordLingo[3] and MyMemory.
The stats here is a whopping 45%. The figure is evenly distributed among all regions. The figure means that nearly half of all translators regularly use a free MT provider.

I will let everyone decide on the figures above, but here are my observations:

- Paid MT versus free MT. 15% translators using paid MT may seem low as of 2016. Note, however, that paid MT is an opt-in (adhering to paid MT is a deliberate act), while free MT is an opt-out (it is active by default, and can be opted out). Many translators run their tool stock. Like most car owners, they rarely open the hood, if ever. Also note that the 15% figure for paid MT usage was under 10% just 18 months ago, which reveals a fast growth rate: it projects into almost 50% by 2020. Free MT, however, remains relatively flat.

- The 15% figure includes an optical illusion, which is typical in statistics, and I will explain it here. Translators upload documents in different formats. Native formats (like DOC, PDF) are markers of an independent translator, dealing direct with clients; while pre-processed formats (XLIFF, TXML, and generally speaking, XML-based formats) indicate that the document was processed ahead of the translator, by a translation agency or corporate translation department. In that case, it is common, almost a rule, that those formats had MT injected at pre-processing time, in which case MT does not appear in the figures above. If high-tech is used ahead of translation, it is likely that artificial translation was used. Well over 75% of pre-processed formats are injected with a mix of Translation Memory and Machine Translation, with MT being more frequent than TM. With pre-processed formats making up nearly one half of the documents today, the real figure of MT use among translators, thus corrected, is above 20%.

9. Conclusion

We can safely say that Machine translation is now mainstream among translators. Concerning fully independent translators, the trend is still modest, but really present, and it is growing fast. As the younger generation steps in, and the emergent economies further develop, that trend can only intensify.

Translation used to be a luxury at prohibitive costs. It is now used at all levels in business and institutions. Two curves are predicting the advent of widepsread use and acceptance of MT at all levels of translation: one is the curve drawn by the need to lower costs in mass translation, the other is the slowly but steady rise in MT quality.

---

[3] WordLingo is not a free MT provider, but it is free for Wordfast users due to a special deal.