Emerging Models in Translation Workflow and New Roles for Translators

Elina Lagoudaki

Abstract

This talk examines the emergence of two new models in translation workflow, as instigated by advances in machine translation, on one hand, and the application of crowdsourcing in translation, on the other.

For the past two and a half decades, the typical translation workflow in companies producing multilingual content and LSPs has revolved around an off-the shelf or customised translation memory system with terminology management facilities, proprietary lexical resources, and a team of selected translation professionals, usually translators, reviewers and project managers, who could access the system and the centralized resources. Remote access to the system and resources has encouraged collaboration among bigger teams of translators working on larger projects and helped ensure consistency, hence quality, among translations. However, even the best equipped and best executed workflows of this type have failed at times to meet particular translation demands, or delivered a satisfactory service after having compromised one or more service quality factors. The difficulties faced currently result mainly from the exponential growth of content that needs to be translated, the shortage of good translators, especially in minority languages, and the cuts in translation/localisation budgets. In order to address these challenges new approaches to translation production had to be invented.

Already, two new translation workflow models are being applied by companies that produce multilingual content: the first benefits from the integration of a customised machine translation system, and the second consists of a translation crowdsourcing and management platform that enables large scale collaboration. The two models can converge. The present study discusses the advantages and issues surrounding the two new models and assesses the roles that translators are invited to play in these new workflows.

For the past two and a half decades, the typical translation workflow in companies producing multilingual content and language service providers has revolved around an off-the shelf or customised translation memory system with terminology management facilities, proprietary lexical resources, and a team of selected translation professionals, usually qualified translators, reviewers and project managers, who could access the system and the centralized resources. Remote access to the system and resources has encouraged collaboration among bigger teams of translators working on larger projects and helped ensure consistency, hence quality, in translations.

However, even the best equipped (with state of the art translation memory systems and top of the line translators) and best executed workflows of this type have failed at times to meet particular translation demands, or delivered a satisfactory service after having compromised one or more service quality factors (for example, by extending the delivery deadlines, overpaying the only available translator(s), employing a non-expert in the subject translator because he was the only one available, skipping a QA round, or increasing the pressure among project members).

It is a common perception among language services buyers and providers alike that current challenges result from the exponential growth of content that needs to be translated, the shortage of good translators, especially in minority languages and rare language combinations, and the cuts in translation and localisation budgets. In order to address these challenges new approaches to translation production had to be invented.

In 2009, SDL International in association with the International Association for Machine Translation (IAMT) and Association for Machine Translation Americas (AMTA) ran a survey to investigate business motivation and utilisation of machine translation as an alternative to traditional human translation, its planned use, and barriers to adoption. Results have shown that there is a much greater awareness and interest in automated translation than there were two years ago and a larger number of organizations are already using or planning to use automated translation in the near future. The survey also revealed that expectations and demands about the quality remain high, with a large number of participants requiring high-quality content as an output of automated translation. Quality is also still considered the main factor in terms of barriers to adoption.

Moving towards MT-centred translation workflows is a reality for many businesses but also for translators who are called in to participate in such a workflow model. Mainstream views held by translators and translation scholars consider machine translation as a threat to the profession because the work that was previously offered to human translators can now be translated automatically by MT systems. This is true especially for certain types of content, for example, technical or for information purposes text; but that does not mean that MT cannot create or has not created other jobs for translators.

Given that more content is being translated now thanks to MT, it is arguable that the specific technology is opening the way to content that was not accessible before, hence increasing the amount of content to be translated. If we add to this the fact that MT is still not perfect, one can instantly foresee the need for human post-editing growing stronger. In effect, translators can be employed as post-editors, after acquiring training on the specific MT system and on basic post-editing principles.

Another set of translation-related jobs that can be created and has been created by MT is those relating to MT data. It is evidenced throughout MT technology literature that the quality of MT output depends on its data and how clean it is. As a consequence, we see the emergence of new services requested from translators, such as data cleaning, assuring data appropriateness, and assuring correct text alignment.

The availability of not-perfect MT output has stimulated the demand of high value translation, according to observations within translation industry circles. Clients shift towards demanding more creative work, stylistically and culturally adapted from their translators. In Europe, the *Transcreation* industry (companies that offer translation and localisation services mixed with creative copywriting) has entered the translation landscape in a dynamic way precisely because of this shift. For translators, this shift cannot be other than positive as they have more opportunities to move from technical uninteresting projects to more creative ones where human touch is indispensable.

Returning to the challenges faced by translation buyers and language service providers, the MT-centred workflow is one way in which some companies have chosen to support and enhance their translation process. Other more technologically adventurous companies, on the client side, have chosen to apply crowdsourcing solutions to their translation sourcing problems.

Crowdsourcing, according to Wikipedia, is a neologistic compound of "crowd" and "outsourcing" for the act of outsourcing tasks, traditionally performed by an employee or contractor to a large group of people or community (a crowd), through an open call. Its principal advantage is that information is aggregated in groups, resulting in decisions that are often better than could have been made by any single member of the group. Crowdsourcing has proven successful in the realm of software development, where successful applications have been created by a crowd of not necessarily qualified software developers.

In the world of translation, the idea of crowdsourcing seems to fit perfectly the void created by the age-old problem of translation quality and its subjective nature. The problem is summarised as follows: Translators and translator trainers know very well that for a given piece of text, with the sole exception of specialised terminology, there are more than one ways to translate it and most of the alternative translations could be considered valid even when the context is clearly defined. If fact, there can be at least 2-3 correct versions of a translation, and this is because every human translator carries and puts into practice his own translation theories and his own personal style. However, at the end of the day the translation to be delivered and the translation to be published must be one. Crowdsourcing seems to give a logical answer to this problem by making use of group in-

telligence: the version that is suggested by most people forming the crowd is put forward as the most suitable translation.

Lately, the idea of translation crowdsourcing has great appeal to content producers and it is proven to be a good way to rapidly expand the language coverage and their global customer base. It can apply to both translation and/or post-editing machine translated content.

One of the most popular examples of successful application of the crowdsourcing model in translation has been the translation of the Facebook user interface. Facebook is to this day the biggest social network in the world, and up until early 2008, it did not offer any localised versions of the site at all. The company managed to jumpstart its international presence with an application called 'Translations', which took the time-consuming and costly task of translating the site and crowdsourced it, asking the network's millions of users to help out (Fig. 1). The process proved to be very efficient: Facebook launched a Spanish site in February 2008, only a few weeks after unveiling the application, and by June it had support for 16 more languages. The French translation of the site was completed in only 24 hours. Facebook is currently available in 75 languages.

Fig. 1: Facebook's 'Translations' application. Spanish translators are invited to help translate Facebook into Spanish.

Facebook users who choose to participate in this task can submit their translation for a number of available English sentences (see Fig. 2) and

their translations are recorded in Facebook's translation memory for sub-
sequent voting. The user has also the opportunity to view translations of
similar phrases which may help him decide on the translation that he will
offer. Additional aid is offered by the availability of some context (e.g. the
location of the English sentence on the user interface) and other helpful
explanation.
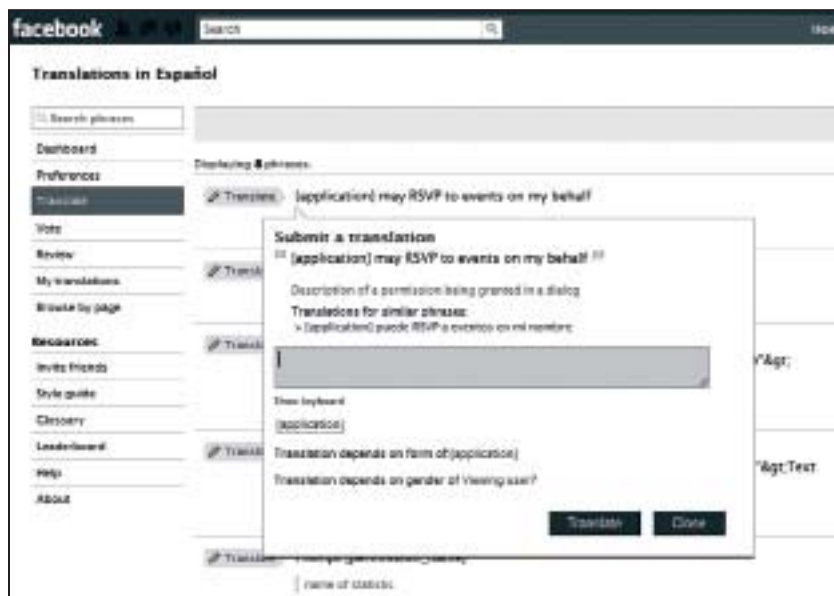
*Fig. 2: How Spanish translators submit
translations for Facebook's English phrases.*

Each and every original English phrase along with its translations that
have been submitted by the users passes onto the voting phase. Here, each
phrase is displayed along with its translations and users are invited to
vote the one they consider as most suitable (see Fig. 3). Once the crowd-
sourcing system is able to suggest a single translation for every English
phrase with confidence, Facebook staff verifies all the translations and
tests everything, making sure that a new localised version of Facebook is
ready to launch.
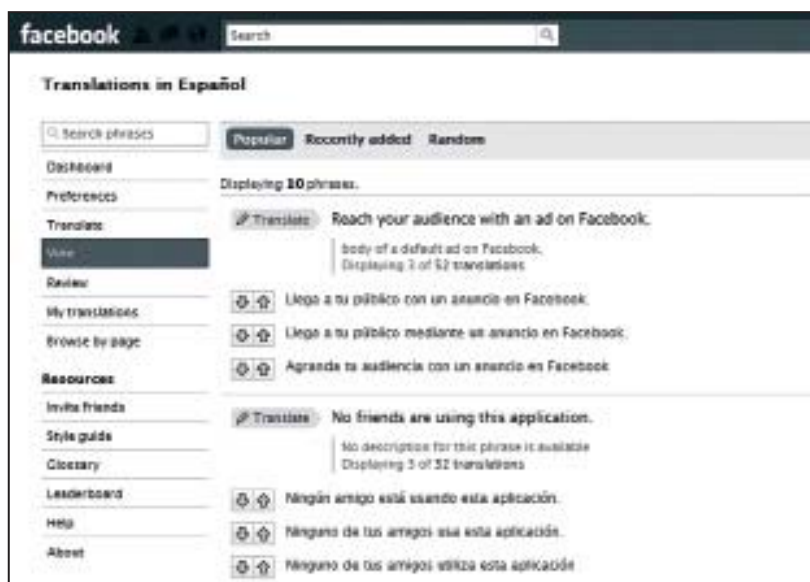
Fig. 3: How Spanish translators vote the most suitable translation.

It is a technique that has worked so well that other social networks (e.g. Hi5, Twitter) have rolled out their own implementations.

Non-commercial projects have also made an excellent use of the translation crowdsourcing model. TED, for example, a non-profit organisation which organises conferences and publishes freely the videos of the conference talks on its website (TED.com), wished at some point to make these talks accessible to all. Translating the talks to all languages of the world via the traditional route of commissioning translation companies was not a viable option, first and foremost because of the massive administrative efforts that would be required and the time it would take for such an enterprise to deliver results. Instead, the TED Open Translation Project launched in May 2009, which allows any talk on TED.com to be translated into any language by volunteers worldwide.

With the help of dotSUB they developed a set of tools that allow participants around the world to translate their favorite talks into their own language (see Fig. 4). To help ensure quality, TED generates an approved, professional English transcript for each talk (this is the transcript upon which all translations are based). Once the talk is translated, TED then requires every translation to be reviewed by a second fluent speaker before publishing it on TED.com. TED controls the final "publish" button.

All translators and reviewers are credited by name for their work. After publication, they provide feedback mechanisms for ongoing discussion or improvement around the translation.

This approach proved to be scalable, and, most importantly, it allows speakers of less-dominant languages an equal opportunity to spread ideas within their communities.

The TED Open Translation Project has translated until now almost 5000 speeches into 70 languages using a pool of over 2000 volunteer translators.
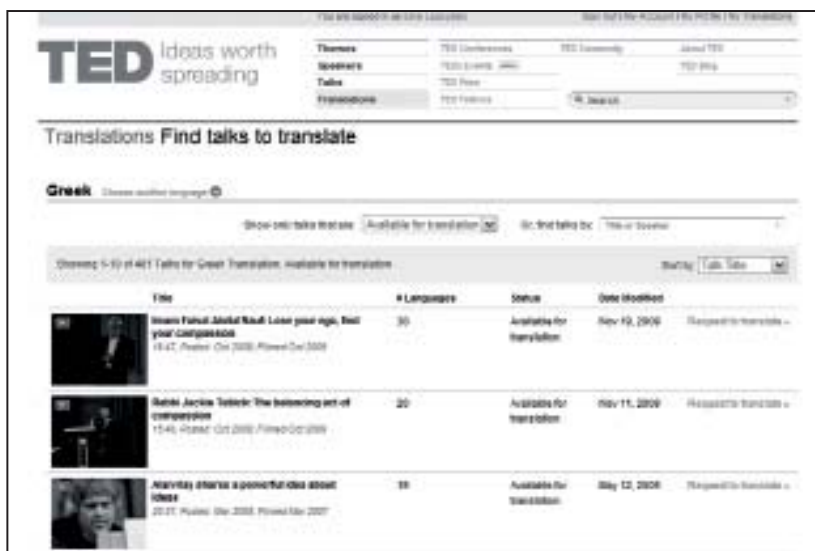

Fig. 4: Translating talks in your language on TED.com

Crowdsourcing has brought many benefits to the organisations which are implementing it, and it is definitely not about saving money (the translation platform, the management of the platform, the management of the volunteers — they all require resources). For social projects, it facilitates the spread of knowledge to all corners of the world. For commercial projects, crowdsourcing is seen as a vehicle to reach totally new markets, to better serve lower margin markets that are currently under-served, and to increase the value of their global brand by further engaging their users.

But crowdsourcing has also been a source of controversy and a subject of a wide debate among professional translators who are worried that crowdsourced translations will push down their rates.

Indeed, the crowdsourcing model in translation is usually based on an open network of volunteers who have bilingual skills, though not necessarily translation qualifications, and a motivation to offer good translations because they are fans of the related product or site. In this type of network, a money-driven professional translator who has no special bond with the product or the site would not stay for long, even if there was some kind of small remuneration. In a less open network of translators (when, for example, participants must meet certain inclusion criteria such as being native speakers of the target language or having a translation qualification), where crowdsourced translations are paid, the rates are defined by the translator's performance; which is in turn indicated by the penalties and premiums given to the work of the particular translator by the other members of the crowd. In such a carefully organised network there are same chances of a translator getting paid less money if he underperforms as of a good translator earning more money. Clearly, a crowdsourcing model has the potential to offer a fair system of remuneration and certainly a fairer than the existing one where the bargaining for the rate is done on the basis of what the translator has done in the past (years of experience, qualifications) and not on how well he performs on the translation project at hand. On the whole, the current model of pricing translation services can be misleading, unfair, it can discourage novices entering the translation profession and can lower the quality of services.

Another concern voiced by translators vis à vis the crowdsourcing model is that translators participating in open crowdsourcing projects loose their power. They as professional translators are not the ones with the authority to suggest and choose correct translations. An amateur's translation may be selected as the best choice because it received most of the votes from the crowd. However, it is high time for translators to realise that they are not always the sole upholsters of translation quality. For example, when passionate users of a product are involved in its localisation it is natural they have naturally better understanding of the product and if they are bilingual they can probably offer pretty good suggestions, possibly better than a professional translator who has never used the product.

A typical translation quality process involves one linguist reviewing another linguist's work. While this is a necessary step, it only captures linguistic opinion. What is missing though is the opinion that matters the most – that from the person buying the product, the end user. By incorporating end-user opinion into the translation cycle, companies can make more informed quality control process and budget decisions by language and by market objective. Even better, user opinion can be used to end the frustration that often surfaces when the reviewer's opinion differs from the translator's opinion.

From a different angle, crowdsourcing can also present unique learning opportunities for translators, especially novices. Translation students can be encouraged to take part in translation crowdsourcing project, perhaps even as part of their curriculum, as

- they would get familiar with the technology enabling collaborative translation;
- they would learn to appreciate the value of collaborating in translation projects;
- they would learn to share knowledge and resources in a collaborative translation environment;
- they would get satisfaction from knowing that they have helped to open up access to knowledge, and finally
- it is a fun way to train translators.

For established professional translators, participating in a crowdsourcing project is a matter of choice. Some choose to participate in initiatives against information poverty (TED Open Translation project example) and bring translation assistance to humanitarian causes (*social translation crowdsourcing*). Others choose to participate in commercial translations crowdsourcing because they consider it a fun way to gain experience in a specialist field that they are interested in and want to specialise in in the future. Others stick to their principles and refuse to participate in any unpaid crowdsourcing projects. In any case, professional translators must be selective when they chose any of these projects to contribute and it is always a matter of having some free time to spend on these or not.