Let's Talk: Trados and the Google Translator Toolkit

By Jost O. Zetzsche

There is no doubt—the situation is tense. Within just weeks of one another, the big-time market leader in translation environment tools (TEnTs) released a completely overhauled version of its Trados tool suite, and the big-time market leader in pretty much everything else released a TEnT of its own, Google Translator Toolkit.

Jeromobot, the patron saint of the modern translator whom I introduced to readers earlier this year,[1] felt there was enough urgency to meet with me and talk things over. Here is a transcript of our conversation.

Jeromobot: Thanks for meeting with me. As the patron saint of the modern translator, I guess I should be cheering translators on, but there are so many things happening right now on the TEnT front that I really needed to talk to someone who has looked at those tools.

> ## The Google Translator Toolkit is an interesting tool and it would be wise to continue watching it, but I think that neither the tool vendors nor translation agencies should be too worried.

Jost: Sure thing.

Jeromobot: So, with Google throwing its hat into the ring, will we all be going to translate.google.com/toolkit and using the Google Translator Toolkit from now on?

Jost: Nope, I really don't think so. The Toolkit is an interesting tool and it would be wise to continue watching it, but I think that neither the tool vendors nor translation agencies should be too worried. But let's start from the beginning. Remember about a year ago when news was leaked to the public about the so-called Google Translation Center?

Jeromobot: Yeah, there was a lot of uncertainty. The likes of ProZ suddenly saw itself potentially going head-to-head with the behemoth Google, and so did many translation agencies.

Jost: Right. ProZ actually launched a

couple of new programs in response to the "Google leak," and there was a lot of chatter among language services providers (LSPs). And rightly so. The focus of the Google Translation Center as it was envisioned back then was translation as well as project and vendor management. Well, as it turns out—whether as a response to all the hoopla that was raised or because, well, because it is Google—they scratched those plans, and what they have now released is totally different.

Jeromobot: So, no management capabilities anymore?

Jost: No, at least not in its current form. Let me give you a quick rundown of what it actually is.

Jeromobot: Thanks, man.

Jost: Presently, the Google Translator Toolkit—I'll just call it GTT from here on out—is free, but there apparently are somewhat vague plans to charge users "whose translations exceed high-volume thresholds." But honestly, I would be surprised if it ever comes to that, because that would mean that professional translators were using it, and, in its current form, that is not who will go for it.

When you open GTT, it presents you with a rather well-designed front-end that allows you to do one of two things: 1) upload an HTML, MS Word 2003, OpenOffice, text, or RTF file; or 2) specify a URL so that GTT can upload the corresponding HTML page. This is a rather meager selection of file formats, if you ask me. When you set the language in the next step, you will quickly notice that there is only one source language you can choose: English. There are admittedly a lot of possibilities for target languages—48 to be exact—but still, that

alone sort of disqualifies it as a professional tool.

Jeromobot: Hmm. Not even Latin as a source language, huh?

Jost: Nor as a target language! Anyway, the next step is sort of interesting. Now you need to choose whether this is a "shared" translation—that is, whether you are using and contributing to a large anonymous translation memory or whether you would like to upload your own translation memory in TMX format. This is the crux of the matter.

Jeromobot: Translation memory sharing?

Jost: Exactly. Not only between you and your co-workers, but also and especially between you and Google. Every time you upload an existing translation memory or perform any translation within GTT, the material will be used by Google for the training of its machine translation engine, even if you declare your translation memory to be "private," and even after you "delete" it. Maybe that is a good thing, but it is sure something folks need to be aware of—and I imagine that your clients would like to know about it as well.

Jeromobot: So what is the actual translation work like?

Jost: Once you have set everything up—by the way, you can also upload a glossary, but only in a really complicated, nonstandard format—your original file is displayed on the left pane in your browser. On the right side you can see a pretranslated version of the file.

The pretranslated material comes from the translation memory or memories first. If nothing is found there—

and even now, several months after launch, the online translation memories do not seem to be particularly well-fed—a machine translation with Google Translate is performed. On the top of your browser you can click on a button—Show Toolkit—which will open a new pane with four different tabs: one for translation memory hits, one for Google Translate's machine translation, one for the glossary, and one for an anonymous dictionary. According to which translation segment you highlight in the Target pane, the different tabs will show the corresponding content—if any is found.

The file is more or less displayed in WYSIWYG (what-you-see-is-what-you-get) format, meaning that it is displayed the way you would see it in a browser or MS Word. Only when you select an individual translation segment is the WYSIWYG view of that segment replaced with a text-only view in which inline codes are displayed as numeric codes with curly brackets—just as other TEnTs do.

Jeromobot: So the actual translation work goes okay, or does it?

Jost: I guess, but it is very mouse-heavy, which is kind of a pet-peeve of mine. Also, many of the features that we take for granted in professional TEnTs—quality assurance, concordance searches, filtering, etc.—are not there. The peer-to-peer quality assurance thing does not work that well either. For instance, a couple months ago, I accidentally added some baloney to the public translation memory, and it is still there—fortunately anonymously.

Jeromobot: So, no saving grace?

Jost: There are a couple of interesting things. For instance, you can ➡

translate a Wikipedia page and upload it right away, or you can cooperate in real-time with someone on a document.

Jeromobot: Okay, so Google Translator Toolkit seems to be more for the occasional translator, but Trados Studio 2009 is a whole different story, huh?

Jost: You could say that! And I do not even know where to start to talk about it. Well, how about a little history?

Jeromobot: I like history!

Jost: Figures. Well, Trados has obviously been around for a while. Twenty-five years to be exact.

Jeromobot: Twenty-five years? I thought you were going to talk *history*?

Jost: Well, are you the patron saint of the *modern* translator or not?

Jeromobot: Okay, okay…

Jost: Anyway, in the past 25 years the requirements for the translator have changed, and Trados has always tried to respond in some way. However, the result was a real hodgepodge of tools that were somehow connected, but you really had to know this "somehow" to make sense of it.

When Trados released version 5 in the early 2000s, they tried to bind everything together in a product called Workspace—and failed badly. The intention was good: to hide all the separate applications and force the user to use a top layer that connected them all, but it did not work because it was immature, users hated it, and Trados quickly withdrew it. It took them a few versions to come up with a new attempt. This time it was called

Synergy. It was a similar idea, but much better built. Agencies liked it, but experienced freelancers still did not. Freelancers were frustrated with early bugs, plus the process seemed to be slightly convoluted for something that could be done more quickly the manual way.

So, this time around SDL tried a much better approach. Rather than working with the piecemeal of the past 25 years, they simply redesigned the application—or I should say, all the applications—to fit into one interface. MS Word as the main translation editor is gone, the database structure of translation memories has been completely overhauled, and all the pieces of the puzzle are accessible from one really good-looking, intuitive interface.

Jeromobot: So this is really good news, right?

Jost: I sure think so. There are some caveats, though. Since Trados Studio cannot do quite everything yet, it comes as a dual version with the old Trados 2007. Certain file formats—PageMaker, Quark, Ventura, and Interleaf—cannot be processed with the new version yet, and alignment also works only through the old version. This has contributed to a bit of confusion with the licensing process. You have to "return" your old license to receive a new one for 2007, plus the one for Trados Studio. While SDL is doing its best to show users how it is

done, I think that pretty much everyone has spent some time scratching their heads about this. There was also some conflict on how long Trados 2007 is valid if you upgrade, but SDL also has been very quick to respond to user concerns in this area.

Jeromobot: Is this a new kind of SDL that we see here?

Jost: Maybe. But I think it also shows a bit of nervousness. See, to switch

We might just see more diversity in the tool market.

from Trados 2007 to Studio is quite a big step. I would say it is quite a bit bigger then the switch from older versions of Office to the revamped Office 2007. Competitors—Across, MemoQ, Déjà Vu, and others—know this and are pointing out that an actual product switch to one of Trados's competitors may be just as easy as learning the new Trados. And they may be right.

Jeromobot: Anything that SDL has to say about that?

Jost: Well, they point to some of their new features that they feel are real differentiators, and several truly are innovative. One is AutoSuggest, in which a bilingual AutoSuggest dictionary from existing translation memories is built to suggest terms interactively as you translate. It is a very clever and effective productivity enhancer, though it is important to keep in mind that you need sufficiently large translation memories for this, and that it does not work for

Jeromobot, the patron saint of the modern translator, and Jost.

Chinese, Japanese, and Korean.

Another differentiator is that PDF files are supported. It is not perfect, but I was surprised to see how well it works, especially on PDFs that originated in MS Word, which is the only format to which you can export the translated file. Imported PDF files from more complex formats such as InDesign require a lot of work on the back-end, and files that come from FrameMaker contain a lot of tags, but this is still the best PDF support for translators on the market yet.

Additional cool features are a real-time and very thorough quality assurance feature, where problems are displayed as you translate. This is a nice preview feature for a number of formats—including, of course, MS Word files—and better support for exchange standards.

Jeromobot: Oh, I like the sound of exchange standards.

Jost: Yes, I do, too. Any file to be translated is actually converted to XLIFF, the translation file exchange format, which has finally become mainstream. The termbase exchange format TBX is also finally supported. I also think that the new TMX format is better than the previous one. One format I am still missing in the new Studio is the segmentation standard SRX, and I am really frustrated with the packaging concept. As in its earlier incarnations—Workspace and Synergy—translation agencies can create packages that are password-protected zip files that create an artificial barrier to prevent the translator from working with other tools.

Jeromobot: Ouch!

Jost: I know! I also do not like the fact that there is no true backward compatibility to Trados 2007.

Jeromobot: But overall your impression is really positive, right?

Jost: Absolutely. It is fascinating to watch the market right now. It is a bit like watching a gun duel in an old Western. The LSPs and freelancers are standing there waiting for the other side to draw first—I mean, to update or not, or to switch to another product.

Jeromobot: So what is going to happen?

Jost: I think we might just see more diversity in the tool market. That would be the ideal scenario, especially when you consider that the new Trados is clearly better than before, there are free and easily accessible tools like GTT for the casual translator, and there are many other great competing tools.

Jeromobot: Right on.

Note
1. Zetzsche, Jost O. "GeekSpeak: The Old is Gone! The New Has Come!" *The ATA Chronicle* (American Translators Association, February 2009), 34.

ata