



GeekSpeak

Jost Zetsche

jzetsche@internationalwriters.com

The Year of the Translator

The GeekSpeak column has two goals: to inform the community about technological advances and at the same time encourage the use and appreciation of technology among translation professionals. Jost is the co-author of *Found in Translation: How Language Shapes Our Lives and Transforms the World*, to be released in October and already available for preorder on Amazon.

I know, I know, it is the Year of the Dragon. In fact, according to the Chinese calendar, I am a dragon myself. But just in the past few weeks I have started to realize that this very well might also become “our” year: the Year of the Translator.

You think it is presumptuous for a technology columnist to make that statement, especially when technology seems to be finding ever more insidious ways to encroach on “our territory”? Well, hear me out.

Recently, I presented a webinar for the Globalization and Localization Association that sketched out the history of translation technology. I am not going to bore you with the entire hour of material, but in a nutshell here is what we found. What started out as technology to support the translator’s process directly (by storing previous translations and giving ongoing suggestions based on that data) morphed into a support system for the complete production chain. By providing features like workflow support, project management, and online accessibility, the technology catered to the needs of translation buyers and multilingual language service providers as well.

At the same time, the development of the core translation features stagnated at its 1990s levels. There were a number of reasons for this development quagmire. First and foremost, we translators just did not show much interest. Understandably, it does not make good business sense for a software developer to focus on developing new features for a group that does not embrace technology—especially while the technology-happy translation buyers and language services providers were clamoring for developers to automate their production processes even more.

This quest for more productive production processes will continue. But a

few things happened in the past year or so that caused developers to sit up and re-evaluate the dated state of their core translation technology. First, machine translation’s renaissance prodded developers into incorporating low-cost machine translation plugins into translation environment tools. At the same time, machine translation developers demanded more nuanced translation memories to build up their data banks. Together, these factors created the perfect conditions for inspiring new enthusiasm among developers interested in bringing our translation technology into the 21st century.

You see, the vast majority of tools previously gave only manual access to data below the level of a complete segment—typically a sentence—even though it had long been obvious that below the sentence is where the true linguistic treasure of translation memories was buried. When statistical machine translation developers (the folks whose machine translation engines are based on large amounts of data) and others called for a new approach to the whole-segment-only approach of translation memories, tool developers responded surprisingly quickly. Today, most major tools have found some sort of subsegmenting technology that provides a way to unearth automatically the real value that translation memories hold and to give us that material at our fingertips.

Now what if these two technologies—access to machine translation and a more granular access to data in translation memories—were actually combinable? What if my tool were to parse automatically my translation memory and offer me all kinds of phrases, terms, and any parts of sentences that match whatever I am currently translating (in addition to the good old perfect and fuzzy matches)? What if it used all those subsegments to give me a first transla-

tion draft, and THEN went out to a machine translation engine and asked it to fill in the blanks?

Or what about the other way around? Sent out for a machine translation but then automatically replaced all parts from the machine translation match that are in my translation memory?

Or what if my terminology database actually contained “intelligent” data, so it knew how to conjugate and decline words, automatically form a correct plural, or use the correct gender as it joins and replaces terms?

Or ... ?

There are plenty of possibilities beyond the ones listed above, and tool developers are already working on and releasing features that take some of them into consideration.

Exciting? I think so. After all, it is unlocking a treasure vault of potentially useful data that we have never before been able to access!

But I am more than just excited. In fact, the reason I think this could play a big part in making 2012 the Year of the Translator is that our fears have been more than allayed.

While many of us worried that technology would push us out, it is actually placing us front and center once again. It is not all about clever workflow solutions and automated processing, though that clearly has its value. Instead, the responsibility for the new “new” thing stands and falls with the translator. Leveraging all the old and new data sources works only with the expertise that we bring to the table. (Of course, this assumes that we will do a better job than last time and engage in essential dialogue with the tool developers).

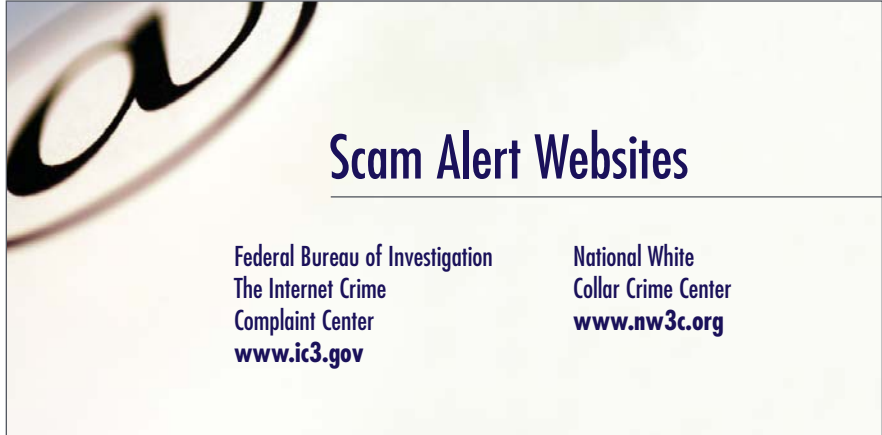
While we are basking in our renewed power, we have the opportunity to step out of the shadows and

Continued on p. 33

GeekSpeak Continued from p. 31

engage with the general public when *Found in Translation: How Language Shapes Our Lives and Transforms the World* is released in October by Penguin. It is high time to claim some of the recognition that our profession so deserves, and I envision this book and the conversations it will inspire to be yet another reason to enjoy this Year of the Translator.

ata



Scam Alert Websites

Federal Bureau of Investigation
The Internet Crime
Complaint Center
www.ic3.gov

National White
Collar Crime Center
www.nw3c.org