



GeekSpeak

Jost Zetzsche

jzetzsche@internationalwriters.com

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 also publishes a free technical newsletter for translators (www.internationalwriters.com/toolkit).

Data Is Good, Right?

Good question. Many novice users of translation environment tools (TEnTs) would argue readily that some data is good, more data is better, and lots of data is best. I remember when I first started using TEnTs. I would eagerly collect all kinds of translation memories that clients sent along for projects and import them into “my” translation memory in the firm hope that this would eventually pay off and give me the leverage for which I was looking. Sometimes that worked. More often it did not.

Today, I would argue that the more data we have at our disposal, the more clearly we see the benefits of quality data versus “lots of data”—for translation memory purposes, that is. You see, if there were indeed so much repetition between projects and clients, good translation would not be so difficult after all, and machine translation, whether of the generic Google kind or the highly trained client-specific kind, would have gained more ground than it has already. So while it is frustrating not to gain the leverage from external data for which we hope, we should also be thankful for it.

Still, one of the reasons I insist on using the term “translation environment tool” rather than “translation memory tool” is that the simple translation memory feature that finds perfect and fuzzy matches is only one of the many relevant features we need to use in these tools. And quite frankly, it is not always the most helpful one (though very welcome when matches from qualified sources are found). If there were no other way to get to data than to hope and pray that the tool finds the hoped-for match (and, again, hope and pray that the match is indeed a good match), then we should indeed stick to very small and highly controlled translation memories. The nice thing is that TEnTs can do much more.

In the past few months, I have written so much about subsegment matching that my editor has started to complain: “Is there nothing else you can write about?” Well, yes, there is, but automatic subsegment matching that tools like Trados, memoQ, Lingotek, and Multitrans now provide has radically extended the usefulness of translation memories, including, and maybe especially, those that come from external sources.

In addition, just think about the terminology extraction provided by default by some tools (such as Similis and Multitrans) and by others as add-on features (such as the SDL line of tools),

or through third-party tools (such as some of the tools by Terminotix). And let’s not forget the manual concordance feature that virtually every TEnT offers (i.e., the ability to search manually for terms or phrases within a translation memory while translating a project), and the self-correcting ability of translation memory matches through the insertion of terminology matches (which Déjà Vu and memoQ offer). Anyway, you get the idea. There is definitely a lot more to translation memories than we originally thought.

ata

ATA Scholarly Monograph Series XV Translation and Cognition

Edited by Gregory M. Shreve and Erik Angelone
John Benjamins Publishing Company

The latest edition in the ATA Scholarly Monograph Series assesses the state of the art in cognitive translation and interpreting studies by examining three important trends: methodological innovation, the evolution of research design, and the continuing integration of translation process research results with the core findings of the cognitive sciences. This timely volume actively demonstrates that a new theoretical and methodological consensus in cognitive translation studies is emerging, promising to improve greatly the quality, verifiability, and generalizability of translation process research.

ATA members receive a 30% discount off the regular price of each hardcover copy in the collection. To order:

www.benjamins.com/cgi-bin/t_seriesview.cgi?series=ATA