

LANGUAGE TECH NEWS



A Publication of the Language Technology Division
of the American Translators Association.

Editor's Note:

In this, my last Editor's Note, I wish to thank each and every one of the contributors to *Language Tech News* since its inception. Your learned, thoughtful articles have gone out to division members and farther afield, helping us to grow and improve in our chosen field of work.

Mention must also be made of our division leadership, beginning with Michael Metzger as first division Administrator, followed ably by Dierk Seeburg at the helm and Naomi Sutcliffe de Moraes as Assistant Administrator. I am greatly indebted to each of you for your guidance, most especially to Naomi for her constant ideas, suggestions, and collaboration

I would also like to thank our proofreader, Barbara Guggemos, for her continued volunteer efforts and attention to detail, which led to each issue going out in near perfect form.

The new editor, Roomy Naqvy, and I collaborated on this issue and he will hold the reigns for future issues. You're being left in most capable hands and I hope you will be as forthcoming with articles as you were with me. I look forward to staying in touch and reading each new issue! □

Lisa Carter
Co-Editor

VOL. 2, NO. 3 / NOVEMBER 2008

IN THIS ISSUE:

Co-Editor's Letter	2
Administrator's Letter	3
Product Review: X-Keys	4
Tool Tips	7
Calendar	8
What is really at stake with MT?	9
Trados Tip	15
Conference Topics	17

49th Annual Conference

We look forward to seeing you at the
ATA annual conference, November 5-
8, 2008, Hilton Hotel at Walt Disney
World Resort, Orlando Florida!



Dear Colleagues:

I have been an ATA member for the last three years and have been co-moderating the LTD mailing list for some time now. About three weeks ago, I received a pleasant e-mail, which offered me the editorship of this newsletter and I have gladly accepted it. So, here I am! I had heard of the ATA around 2000, but I found the fees forbidding and took time to decide before becoming a member. I was a ProZ Moderator for English<>Gujarati and their Literature Forum from 2000 to 2007. I asked a number of translators if the ATA membership was worth it. I recollect people telling me it wasn't! Now I know from personal experience how wrong they were.

I hope I will be an interesting mix for you. I come from India, where I live and work. I teach English literature and grammar at a University in New Delhi, the Indian capital. I began working in 1995. I was responsible for introducing the Diploma in Translation Proficiency at our university. I also teach English and Public Relations to Media students. I do a number of things and like being versatile in life. So, I do translate and I do translate a lot. When I look at how a full-time translator is defined, in terms of hours worked per week, I realize I shouldn't call myself a part-time freelancer!

As well, I do have several hobbies and I like maintaining my blog. But let's keep it short for now. You can always look me up on my ProZ or ATA profiles. Or e-mail me!

I love communicating with people. I just love it. I am really excited about editing this newsletter. I would like to express my gratitude to Naomi and Dierk for getting me on board. But without Lisa, the great Lisa Carter, this issue wouldn't have been published. She has been great and really helpful. I hope I can trouble her sometimes even when she's not formally called The Editor. ☐

Roomy Naqvy
Co-Editor

LTD Listserv Co-Moderators

We are pleased to announce (and thank!) the new co-moderators of our listserv:

**Katrin Rippel (EN>GE translator) and
Roomy Naqvy (EN<>Gujarati, Hindi)**

A round of applause, please!

Language Tech News

Vol. 2, No. 3 November 2008

Copyright©2008

American Translators Association

225 Reinekers Lane, Suite 590
Alexandria, VA 22314
Telephone (703) 683-6100
Fax (703) 683-6122
ata@atanet.org
www.atanet.org

Editor

Lisa Carter

lisa@intralingo.com

Editor

Roomy Naqvy

roomynaqvy@gmail.com

Proofreaders:

Barbara Guggemos

Contributors to this Issue:

Tuomas Kostianen

Naomi J. Sutcliffe de Moraes

Dierk Seeburg

Sam Stallard

Rosana Wolochwianski

Layout:

Cindy Gresham

illustrator@cgrisham.com

LTD is the

Language Technology Division
of the American Translators
Association

LTD Administrator:

Dierk Seeburg

info@icotext.com

LTD Assistant Administrator:

Naomi J. Sutcliffe de Moraes
www.justrightcommunications.com



From the Administrator

Welcome to the fall edition of *Language Tech News*,



the LTD newsletter! By the time you read this, the ATA annual conference will be nearly underway and preparations will be in high gear. For the first time this year, we have organized a social event for LTD members and anyone else interested in it. Please join us at the LTD Wine & Cheese Networking Event on Friday, November 7, from 5:30-6:30 p.m. We hope that you will join us for this reception and network with members of the Language Technology Division while enjoying cheese, bread, and wine. Please see the Conference Registration Form <www.atanet.org/conf/2008/register.htm> on how to sign up and get a ticket for this event.

I hope you will come to the Language Technology Division Annual Meeting, Conference Code: LT-6 <www.atanet.org/conf/2008/byspecial.htm#LT-6>, which is scheduled for Friday, November 7, from 4:00-5:00 p.m. This is your opportunity to get involved and I encourage you to join your colleagues and me in shaping the division's future.

Language Technology made the news again in the last three months in several areas. Arguably the most notable event was search engine giant Google's introduction of their Google Translation Center <<https://www.google.com/accounts/ServiceLogin?service=gtrans>>. Now in private beta, this offering looks to be a platform for connecting translation buyers and translation service providers while possibly being linked to Google's own statistical machine translation service, Google Translate <<http://translate.google.com>>. Google seems to have picked up on the notion that there are opportunities in translation management and automation. We'll be keeping our eyes and ears open on this one.

Speaking of machine translation, the TAUS Data Association <www.translationautoma

[tion.com/best-practices/solid-foundation-for-taus-data-association.html](http://www.translationautoma.com/best-practices/solid-foundation-for-taus-data-association.html)> was incorporated with 47 members at the beginning of September. It aims to "establish a secure and legitimate platform for storing, sharing and leveraging language data; access to large volumes of trusted language data for increased translation efficiency and quality; industry collaboration to harmonize multilingual terminology." Questions have been raised <www.globalwatchtower.com/2008/06/26/taus-tda-charter> about the organization's goals, among them the return on investment for translation service providers and translation buyers as well as usefulness and quality control of the available translation memories.

This marks our current newsletter editor Lisa Carter's last edition. In the name of our membership I would like to express my sincere appreciation for a job well done. Under her stewardship the newsletter has grown into its own, thanks to her high standards of quality and attention to detail. We wish her all the best in her future endeavors—thank you Lisa!

We are happy to announce a successor in colleague Roomy Naqvy, who will be taking over the editorship from Lisa. She is kindly guiding him through the transition with this issue. Roomy hails from New Delhi, where besides being a translator he teaches English and a special course on translation as Assistant Professor at Jamia Millia Islamia University New Delhi. Please help me welcome Roomy to his new position and let him know any questions and suggestions you may have. □

See you in Orlando in November!

Dierk Seeburg
LTD Division Administrator

X-Keys: Convenience to the X-treme

By Sam Stallard

Having fallen victim to a repetitive strain injury and

struggling to maintain my livelihood, I came across the X-keys line of products by PI Engineering (www.xkeys.com/xkeys.php). The



Figure 1



Figure 2



Figure 3

company offers a line of personalized shortcuts keypads, which are small peripheral USB or PS2 consoles with banks of easily programmable push-buttons.

I ventured to purchase a copy of the X-keys Stick (Figure 1), the smallest product in the line, to check it out—afraid of wasting my money on something too complicated or superfluous—and was quite pleasantly surprised. I found the thing so convenient that I soon bought another one, the larger Desktop model with 20 buttons (Figure 2), and then recently sprang for the 58-button Professional model (Figure 3).

These currently list for \$100, \$120, and \$170 respectively. There appears to be virtually no dealer availability and no better prices, especially in Europe.

In short, I highly recommend these products

for translators. For those who suffer from or are concerned about RSI, this is nearly a no-brainer, but even the more fortunate individuals in our industry will likely find the convenience to be worth the investment.

Though the manufacturer obviously has a sense of humor as “The No Slogan Company,” they may not be fully aware of the relevance of their products to translators in particular; we appear to fall into the “other” category for whom the products are claimed to be ideal. This potential slight can be overlooked, however, in view of the benefits the products afford!

Whatever efficiency strategies you may be using, from simple keyboard shortcuts to macros to CAT tools to regular expressions and dictation software, X-keys makes executing them faster and more...I have to say it again...convenient.

Many translators will also be familiar with the problem I eventually encountered that while a few gamer keyboards offer a handful of customizable shortcut keys, this is not enough for sophisticated usage. Even if many boards nowadays at least have non-customizable cut-and-paste hotkeys, most of these are just multimedia buttons like volume adjustment and so on, or even more useless, like ‘search for file.’

After customizing a few Word commands and maybe some program startup shortcuts by assigning frequently used functions to, for example, the function keys F1-12, one starts to run out of keyboard shortcuts without stepping on useful existing shortcuts. Word shortcuts involving the ALT key are especially impacted, as well as CAT-tool shortcuts. Eventually you end up with four-key combinations like CTRL-ALT-SHFT-T (the ‘Vulcan Death Grip’) that call into question the idea of a shortcut. And you start losing track of what commands have been assigned where. Again, here is where X-Keys comes in, giving you one-key execution and keeping you organized.

Some real-world usage examples (nitty-gritty stuff):

Say you work with Trados and have assigned a delete-to-period macro to CTRL-ALT

D or F12. Program an X-key to execute F12, then ALT-NUM+ for Set/Close Next Open/Get. This turns four or five button pushes into one and might perhaps be used a hundred times a day.

Or, building upon this concept, consider for example that when you use a delete-to-period macro, you often first have to move to the end of the word you want to delete from, generally by using CTRL ->. This puts you at the start of the first word you want to delete from, leaving one space between what is going

to be the last word in the sentence and the period, like this . So, you first either have to use backspace every time, or write a second macro with a backspace in it, or take out the spaces all at once using another macro. But what if you are working in a non-Word environment like Tag

Editor where Word macros don't work? For a larger file of say 14,000 words, I am guessing this could mean having to hit that backspace key roughly a thousand times. Plus, in the specific scenario the original delete-to-period macro itself won't work. So, as a workaround you can, in no time at all, program one or two X-keys to first backspace, then delete the next ten or so words via repeated CTRL-DEL (generally enough to take you to the end of a sentence—segments remain unaffected by excess deleting in Tag Editor), and then execute ALT-END + ALT-HOME (sticking with the example of TRADOS) to close the segment and open the next one. This can save a world of hassly button pushing.

The more frequently used, the greater the utility, but benefits can be had with less complicated yet frequent functions such as (among my personal favorites): ALT TAB, PASTE SPECIAL—UNFORMATTED (after recording a macro with a keyboard shortcut)*, SPACEBAR CTRL V (paste) and CTRL V+ENTER, i.e. for Google searches. Other handy shortcuts include 'copy next word', 'replace next word' (CTRL-SHFT ->, CTRL C, CTRL H, CTRL V, TAB)—also modifiable for PowerPoint—

'replace all' (from Replace window — CTRL V, ALT A, ENTER, ENTER, ESC), CAT tool combination commands such as 'concordance + next word' and 'cleanup + open file.' And that's just a sample—the possibilities are manifold!

If you use regular expressions, you probably get tired of typing things like ([0-9]{1,2}).([0-9]{1,2}) (M??. EUR). Instead, just assign that text sequence to an X-keys key and it becomes a one-button-push affair. You could even have the Replace window open for you automatically by adding CTRL H to the sequence in front, and TAB at the end taking you to the Replace With line, with or without additional pre-programmed entries.

Another usage: the leading dictation software is an excellent product—especially for RSI sufferers, it is a lifeline—but it has its well-known limitations. For one, not everyone has the latest cutting-edge processor to obtain the best speeds, and common voice-executed commands can be slow and are often incorrectly recognized, even for experienced users. These issues can affect even those few users who elect to purchase the expensive Professional version with advanced voice macro capability.

Instead of having to wait around for the software to properly recognize such frequent sequences as 'press home' + 'paste that', you have a one-button alternative once programmed into an X-keys device. Another big one is Dragon's only sporadic ability to properly understand instructions to capitalize a word, so you frequently have to go back to the start of your sentence, delete the first incorrect word and recapitalize the following word. This was particularly a problem for me in Tag Editor. Solution: Record a macro to capitalize the next character, then program a button to press HOME + CTRL DEL + the key combination for that capitalization macro! In summary, one can do a lot this way to compensate for the deficits of dictation software.

The button action of the devices is soft and pleasant, the key caps are easy to remove for

You will find the convenience to be worth the investment.

* (see <http://pubs.logicalexpressions.com/pub0009/LPMArticle.asp?ID=128> for reference)

inserting labels and don't break or chip easily from frequent relabeling. Button label design and printing is also uncomplicated using a separate, included module. Some colorful pre-printed labels are provided, but these are mostly for functions that will be irrelevant to translators.

Just assign a text sequence to an X-keys key and it becomes a one-button-push affair.

It seems all their product drivers are on the one disk provided, so when I was loading them I installed a bunch I didn't mean to because I was confused. According to the website there is now improved support for Vista. The operating software is simple enough to figure out, but the instructions are rather spare. The learning curve is still quite manageable. The Help Topics under the Help menu provide a sort of operational manual to consult if you are baffled. But if you really can't get running, give them a call. I bet they'll help you, especially if you tell them their product was recommended in an ATA publication for translators. They have upgraded the macro-writing software used with the devices and probably have improved on a couple of minor inconveniences, like non-portability of key encoding if you decide to move a command sequence to a different key.

To set up the keys, you just open the programming interface Macro Maker, select the X-keys device to program, select the key to program and go to Create/Edit macro (F2). After naming the operation/key label, you hit Start and then just hit whatever key sequence you want. Then hit OK, Save, and Assign to Device. The AutoRepeat option is crucial to remember for some functions (hold down button and command repeats), and the Insert Pause option can also be useful at times when you are automating a function that takes a second or so and gets messed up if it hasn't finished.

Macro Maker is accessed through the overarching module Macro Manager, which usually loads itself to the bottom right tray in Windows XP, although sometimes it doesn't and has to be started manually. Macro Manager is useful when you are using more

than one device and/or more than one programming layout; it is not of primary interest. All this may also have in fact changed with the new interface, which I have not adopted because I don't have the time to reprogram everything.

Each device is layered, giving you twice the number of assignable functions as there are physical keys. That is to say, for a given button you can program two functions. It is best to have all the most frequently used commands sequences on one layer (green) and less frequently used ones on the other layer to reduce the amount of toggling back and forth from layer to layer. The toggling between layers function has to be assigned first. Keep in mind if you have your keyboard set to toggle between languages, your commands may not work or work properly unless you have the keyboard set to English if you programmed them from that setting.

I was originally attracted to the intriguing-looking X-keys Stick, but have found it less ideal in terms of placement, i.e., generally above or below your keyboard. The 20-key square Desktop model, which goes next to your keyboard, may be best if you want to check the concept out first before investing more money. However, I can imagine any translator would have no problem utilizing all the additional (58x2) keys provided by the Professional model right off the bat.

Besides dictation software, this is probably the next most important tool for RSI sufferers, and for everybody else a nifty productivity enhancer. Sliced bread anyone?

Sam Stallard, M.A. German, is a state-certified translator in Germany. He can be contacted at deutsch-englisch@online.de.

Share Your Knowledge

Language Tech News is a great way to get your name out to your colleagues, to share your expertise, and to give back to your division and your association. Send in an article and share the wealth of knowledge you have!

Tool Tips

by Naomi J. Sutcliffe de Moraes

More screen space

In this issue, my tip is related to hardware. The last time I purchased a new desktop computer, I noticed that there was an option to purchase a video card with two input connectors, which would allow me to attach two monitors to my PC. I had two monitors in my office, but the smaller, older one was in the closet. I took it out, dusted it off, and voilà! The result can be seen in Figure 1. For those of you not familiar with this kind of setup, the second screen is not a copy of the first, but rather an extension of it. The mouse pointer glides continuously from one screen to the other, and you can drag things



Figure 1 – Two monitors on one PC

around as if it were one big screen. Windows has this capability through what it calls Dualview. Linux also has this capability through the video driver. (See references at the end of this tip.) No special software is needed, just a video card for two monitors or two or more video cards (usually of the same type, for compatibility). A few years ago, I thought “if two is good, three would be great!” and I bought another video card. Unfortunately, my PC’s power source was not powerful enough to run the additional video card. Even worse, the power source was proprietary and I could not simply replace it with a more powerful one. Be warned! If you are buying a new computer, you can ask the person setting it up to make it work for as many monitors as you might want. Look at

the references at the end of this article to learn more about how this works, especially if you want to alter an existing PC.

As an aside, in case you are wondering, I have my monitors up on a platform and my keyboard under it because my physical therapist told me to have my entire forearms resting on the table when I type, not just my wrists, and it really is quite comfortable typing that way.

Another option is to use your laptop monitor as the principal monitor and plug an extra monitor into it. If you activate Dualview in Windows, the second monitor will act as an extension of the laptop monitor rather than a copy. This also works in Linux, as far as I can tell.

Using a second PC’s monitor

Do you have a spare laptop hanging around your office? Use its screen as an extra monitor! See Figure 2 for an example. The software program MaxiVista (\$29.95, with free trial version) can allow you to use one or more laptop screens as extra monitors. They must be connected to the same network as your principal PC. MaxiVista works only with PCs running Windows.

Control two PCs with one keyboard and one mouse

The same program, MaxiVista, will let you access the desktop of a second PC (connected through a network)



Figure 2 – One PC with two monitors plus one laptop

without switching to the keyboard/mouse of the second PC. You just slide your cursor over to the other monitor and the second PC receives your commands. You can even cut and paste text, files, and folders from one PC to another. What a great idea! Definitely worth thirty bucks.

If you are updating a current system to use multiple monitors, www.multiplemonitors.org supplies some information on hardware compatibility. Your safest route is to use identical video cards and thus identical video drivers. Avoid mixing PCI and AGP-type cards. I have been very happy with my Nvidia hardware and the accompanying nView Desktop Manager. □

References:

Microsoft Windows built-in multiple-monitor capabilities:
http://www.microsoft.com/windowsxp/using/setup/learnmore/northrup_multimon.msp

Debian Linux built-in multiple-monitor capabilities:
http://aufrecht.org/blog/one-entry?entry_id=23251

MaxiVista (to connect two PCs): <http://www.maxivista.com>

Multitple Monitors.org: <http://www.multiplemonitors.org>

Naomi J. Sutcliffe de Moraes is a Portuguese into English translator. She has a PhD in Linguistics from the University of São Paulo, but prior to becoming a translator she worked as an engineer and physicist. Her specialties are medicine, engineering, and law. She can be reached through www.JustRightCommunications.com.

Calendar of Language Technology-Related Events

For more information, go to <http://atanet.org:80/calendar/>

DATE	TITLE
October 25	<i>Assn of Translators & Interpreters in the San Diego Area (ATISDA)</i> Wordfast Workshop San Diego, CA
October 29–31	<i>Languages and the Media</i> 7th International Conference and Exhibition Language Transfer in Audiovisual Media Berlin, Germany
November 22	<i>Northern California Translators Association (NCTA)</i> Trados for Beginners San Francisco, CA

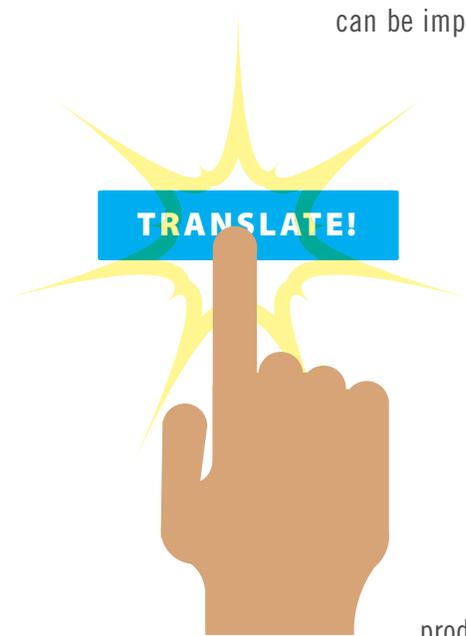
■ What is Really at Stake with Machine Translation?

An Overview of its Impact on the Different Stakeholders

by Rosana Wolochwianski

Machine Translation (MT) gives rise to all kinds of

reactions and concerns. Large companies, always eager to reduce costs, wonder how it can be implemented and if it is worth



the investment. Final users have divided opinions: some of them enjoy the possibility of being able to access a

low-cost or free translation, even if it is imperfect; others express disappointment because of the (lack of) quality of the results.

Translators seem to be worried at the possibility that these programs might displace them in the production chain, turning them into mere editors of pre-

translated material. They wonder to what extent this development will affect work availability and compensation. MT researchers and developers have been asking themselves time and again, for over 50 years now: why can't these programs deliver better results?

Machine Translation vs. Human Translation

For translators who deal with language on a day-to-day basis, it is quite obvious that the quality of these programs is still poor in many cases, and this generates jokes and anecdotes among us, making us feel that we are all-powerful and cannot be replaced. To reach publishable quality, MT output requires human intervention, either in the authoring phase (by means of controlled-language efforts) or at a later phase, through post-editing. There is no

doubt that MT entails quality constraints. However, it is evident that the scope the democratization of technology has reached in this globalized era generates consequences that we would have never imagined. It is estimated that today there are nearly 1.5 billion people online, and almost 104 million web domains actively in use (which account for about 30 billion pages, according to Netcraft).

The amount of information that circulates today is so huge, and the eagerness to access it is so urgent, that it is almost absurd to think that only a group of qualified professionals producing flawless 300 to 400 hundred words per hour can satisfy such a great demand. Human Translation entails time and volume constraints. So, it is necessary to admit that a good part of this huge amount of information—especially information that, in any case, would never reach the hands of professional translators, either for lack of time or lack of budget—will possibly be processed by MT programs. An alternative to zero translation has emerged.

Dissemination vs. Assimilation

MT programs can normally be used in two directions. In the original development plans for MT, the objective was to create a tool which would be capable of translating text for distribution by the original authors, in a process traditionally called “dissemination” of information towards the users. So far, these types of commercial systems (e.g., Systran ProfessionalTM, Language Weaver SMTS, SDL Enterprise Translation ServerTM, and so on) render a result that needs to be reviewed and corrected by human translators in order to achieve an acceptable level of quality. They have been widely criticized by the translation

community. However, there is another reality to which we need to pay attention. Everyday, millions of people click on the links to the free MT tool on an Internet page (e.g. Yahoo!, BabelFish, SDL FreeTranslation, Google Translate, and so on) to perform searches, translate short messages, etc. This merely approximate translation that is retrieved by the final user with the intention of roughly understanding the central idea of a text is what we call ‘gisting’; in a way, it is opposite to the “dissemination” process, and we can

call it an “assimilation” process. This has been an unexpected result of the extended and decentralized use of the Internet.

In this sense, it is necessary to accept that translation is no longer exclusively a translator’s job, with a traditional translator devoting hours to finding the best possible translation for a word. When we think of translation today, we must

also think of a cybernaut trying to find a resource online and making use of the free MT programs available to get to know, at least vaguely, what a page is about, or how to fix a bug in a computer program. Additionally, we must think of a tourist trying to decide which hotel or which meal would suit his needs best during a trip. So this new complex reality, flooded with information and urgency, can be considered from different viewpoints, which reflect a variety of interests.

The Different Players and Their Points of View

When it comes to assessing the usefulness of MT, we come across the pragmatic perspective of users (the cybernaut, the tourist, someone in a chat room), who focus on what works for them, and might not care much about quality—at least when it comes from a free resource and provides an instant solution to immediate needs. Their questions are more

like, “Does this technology work?” “Does it solve my problem?”

Researchers, by contrast, have an academic perspective. Curious and perfectionist by nature, they are not satisfied with the lack of quality of the results. They want to understand why this technology does not work better and, possibly, find a solution to improve its performance.

Let us not forget the large corporations with translation departments and the translation companies that need to translate endless manuals or support documentation into several languages in a much reduced time, and at the lowest possible cost—in other words, companies that need to replace labor with technological resources as much as possible. These companies want to know if MT’s return-on-investment is justified, and if the quality of the products they deliver will be compromised.

Last but not least, what about professional translators? Many of them are reluctant to admit the usefulness of MT, as they see it as a tool that, together with translation automation processes in general, has the purpose of displacing them in the production chain and confining them to the role of “editor” of material spit out by MT, instead of letting them be the one who translates texts from scratch. This might be true for certain types of texts, and it is already the case when Translation Memories (TMs) are used. It becomes, no doubt, a bit alienating for translators, who probably dreamt of translating classic literature in their school days. It does call for a redefinition of how translators are compensated. It could signal a shift from a per-word compensation to a per-hour compensation scheme, or to other new alternatives which could be explored further

Different MT Systems

For almost 50 years, MT research focused on what is known as rule-based machine translation (RBMT). This is the classic system, represented by Systran and all its successors. It is the type of MT program that is commercially more affordable or more accessible

It is almost absurd to think that only a group of qualified professionals producing flawless 300 to 400 words per hour can satisfy such a great demand for information

through the Internet. It makes use of bilingual dictionaries and a set of lexical, syntactical and semantic rules for each language pair. These MT systems are “black box” systems, and it is very costly and complex to develop them for every new language pair. Other older MT systems are laboratory projects linked to Artificial Intelligence (AI). These systems aim to introduce into an MT system enough knowledge of the world in order to make it “think” and interpret the way a human being does. They are usually university-based projects (KANT, UNITRAN, Carnegie Mellon).

*Everyone in the industry
(and not only translators)
should understand that
the meaning of an expres-
sion does not exist beyond
the usage it is given
in a certain context.*

The 1990s saw new alternatives being explored. After almost two decades of TM usage, a huge amount of aligned bilingual material was available. The new challenge was: could this corpus of aligned material be used to feed an MT system and, combined with good search engines, could it produce an MT

program that was capable of learning through successive translations, and that could be easily expanded to new language pairs? That’s where statistics-based (SBMT) and example-based (EBMT) systems entered the scene, with their probability-driven and pattern-driven approaches, respectively. On the commercial level, SBMT was initially offered by Language Weaver, and is now also offered by Asia Online.

There is even a project of context-based machine translation (CBMT) supported by a company called Meaningful Machines, which claims it uses bilingual dictionaries and monolingual extensive corpora in order to train its prototype MT system. The new automatization efforts are also directed at building hybrid systems by combining the best of all worlds: RBMT + SBMT + TM, etc.

The Limits of MT Systems

More than fifty years of research, different technologies applied, new investment, lots of previously aligned information fed into the

systems, and the MT systems still do not work that well; they are still criticized and frequently the object of jokes and anecdotes. Why? Well, the preliminary answer is quite simple: because translation depends on a unique human capacity—that of interpreting meaning, making inferences and conveying sense. Pragmatic processes allow us to close the gap between the semantic representation of a given text and its interpretation as a statement realized within a certain context. What is said does not just consist of the conventional meanings, but is also the result of reference allocations, disambiguation and the enrichment of some expressions—which takes us from the level of conventional meaning, to that of communication.

The result any MT program can produce is just a proposed target language equivalent, the result of rule application and/or matching efforts, not a translation in its proper sense. Everyone in the industry (and not only translators) should understand that the meaning of an expression does not exist beyond the usage it is given in a certain context, that there is no preexisting translation a program can just find, deduce or decode. On the contrary, a piece of translation needs to be figured out on the spot; it is not a mere transfer of meaning from one language to another. That is precisely why different translators can produce different translations for the same original text, and why the same original text can require different translations in different contexts. So retrieving an exact equivalent from a database is just not good enough. Such a process relies on extra-linguistic human knowledge such as culture, experience, beliefs, assumptions, and above all, interpretation skills and common sense, something machines do not have and will most probably never have.

New Models: The End of the Utopian Phase

So, at this point we might be tempted to think we can relax. Due to the very nature of human language, MT programs can’t translate the way human beings do. If program developers do not admit this, their attitude can

certainly be viewed as a form of voluntarism. This is not a question of time. It would not happen in another five years. The issue is that the industry has reached a very similar conclusion, and has decided to change its perspective to a more realistic one. It has concluded that the classic idea of “Fully Automatic High Quality Translation”

(FAHQT) is yet to be developed, but that a form of “Fully Automatic Usable Translation” (FAUT) (as an alternative to zero translation) can already be achieved and leveraged. This 180° shift in perspective has revolutionized the translation industry at many levels.

For the last fifty years, researchers have been struggling to create tools that can translate with the same level of quality a human being does. This, obviously, has not been possible to date. However, the need to access translations in huge volumes and almost in real time is so great and so urgent that, in many cases, the user does not care about quality. There are situations in which, due to lack of time

or budget, an imperfect translation is preferable to having no translation at all. “Let’s embrace the imperfection of MT,” claims the Translation Automation User Society (TAUS)—founded in 2005—in its vision statement.

TAUS champions a new localization model in which the final users and the market are the ones that control the translation flow, not the publishers. They are working on the “self-

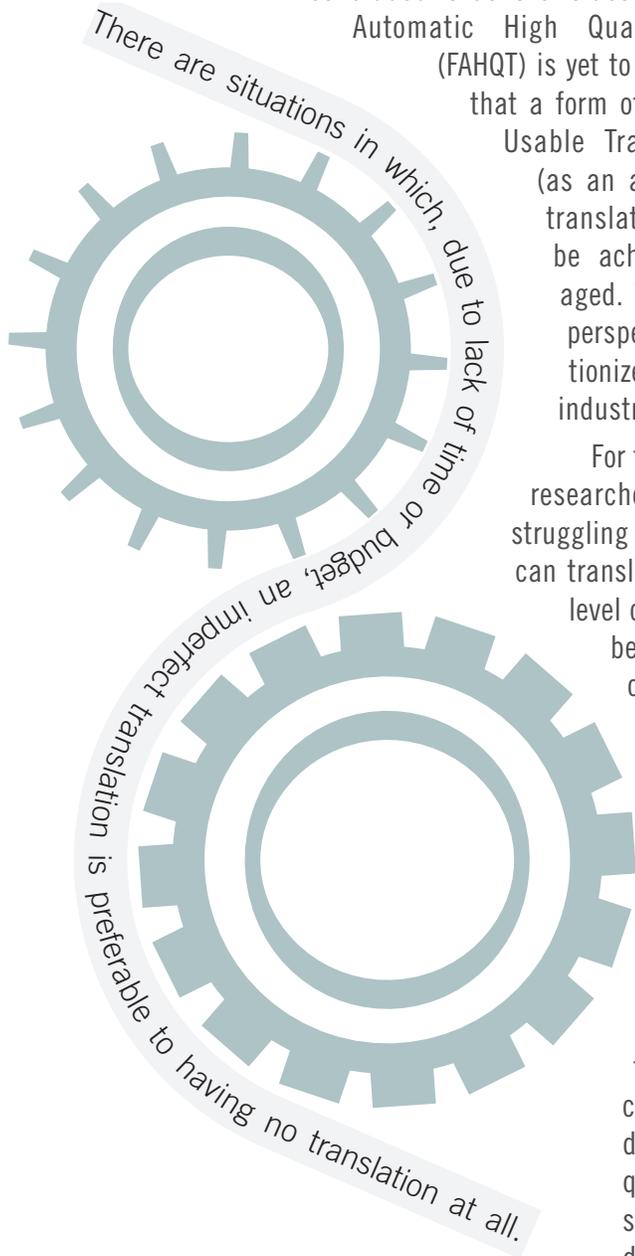
service” information model propagated by Google, and this is becoming a reality, so we should be aware of it and act ourselves with common sense, by accepting this new reality. TM sharing and the development of large TM repositories are under way: the TAUS Data Association was incorporated this year by 40 founding members, with the aim of selecting and pooling data to increase translation efficiency and improve translation quality.

Of course, that makes us wonder: Could a potentially larger TM repository, even if organized by industry domains, be effective in order to feed and train MT systems, given the subtle context restrictions any piece of translation poses? Anyone who might have tried to merge TMs from different clients probably knows the matches rendered are usually far from relevant. Mr. Yves Champollion [2002] warns us in one of his articles against this making up for lack of relevance with size, with the use of “blind, random TMs.”

MT in Practice

Notwithstanding all that was said, one thing remains true: technology is neither good nor bad. It is just a tool, and it all depends on who uses it, how, and with what purposes. I have read and heard stories in which MT is applied creatively and effectively in order to reduce lead times, cut costs, facilitate searches, preselect materials to be translated, and even please translators with new negotiated win-win compensation schemes.

I will not discuss these examples due to lack of space, but they are out there, and we have all started to hear about them: MT applications for the translation of knowledgebases / customer support (translation on demand / prioritizing localization needs); MTM solutions: TM + MT combined in high-volume time-restricted projects; automated translation of intranets and news bulletins for multilingual employee bases (for the sake of reaching out / keeping confidentiality); patent search engine translation projects (like the one of the European Patent Office, based on the Japanese counterpart example); translation of extranets (like the case of movies distributors, as



well as product catalogues); virus alerts (where instantaneity becomes a must); and so on.

Some Critical Points

Where are we then? There are plenty of scenarios in which MT can be applied, either for a less-than-perfect-quality translation, or for a pre-translation to be polished by human professionals at a later stage.

However, there are still plenty of concerns. For example, how can the quality of MT be measured? So far there is one predominant standard called BLUE (Bilingual Evaluation Understudy) by IBM, which identifies in MT output similarities to a reference human translation, and there is also a painstaking process metric carried out by humans. Another concern is: how can return-on-investment be justified without clear metrics? If a company has to assess the cost of controlled language + post editing + SBMT continuous training, is investment in MT still profitable?

On a different note, how are clients' expectations handled? Are translation companies conscious of what they sell when they offer MT? If they post-edit everything,

there might merely be a risk of loss of profit. However, if they offer it directly to clients as a low cost option, are the clients aware of the kind of quality they will receive? As it is popularly said, there is no second chance at making a good first impression. Couldn't MT become a business boomerang if the client is disappointed? Finally, how can the resistance of translators be handled? Are there new balanced options as regards productivity per hour which could entice professionals into working with MT?

Post-Editing: A New Job Opportunity?

It is clear that MT does not pose a risk of lack of work for translators, at least not at the moment. Its use is restricted to certain highly

repetitive areas, integrated in the job workflow as just another tool. There will still be a need for translators for many other areas in which quality is non-negotiable, like marketing, law, literature, etc. So, well-seasoned translators will still find the way to go on working without using MT if they want. In that case, who will work on MT post-editing?

A few years ago I came across an article in *The ATA Chronicle* in which post-editing was considered a new job opportunity [Schwalbach & Zearo, 2006]. There the authors explained that post-editing is a type of work with its own characteristics, for which we can get specifically prepared by developing special skills like speed, understanding of the different post-editing requests (complete, minimal, partial), and so on. I wonder if we can really tell editors to do a "partial" editing. I also wonder how they feel when they have to edit terminology but ignore grammatical mistakes, or vice versa. The article also highlights that the best candidates for this type of work are, of course, the newcomers, the junior translators, as they are more open-minded and they need the work. This led me to some reflections.

A French anthropologist called Marc Auge [2000] points out that it is at the moment we develop our writing abilities, that we discover the subtleties of reading. We can all agree that this is certainly true. When we learn how to read, we do not get hung up on the differences between an "s," a "c," or a "z," between "v" and "b." We just go on reading. But it is when we intend to write, to produce, that we start to have doubts about "which is the correct letter to use here?", and we become aware of the subtleties of language.

I think there is a possible analogy with translation work here. We learn to translate by translating. It is by deciding creatively, each and every time, and by making mistakes time and again, that we become well-seasoned translators, and acquire that subtlety that makes us good translators. So, if a new translator enters the industry as an editor of material which has been preprocessed by an automatic program, will he really be able to

Rosana Wolochwianski is an ATA-certified translator from Rosario, Argentina, who has been in the industry for over 15 years.

She is the Founder and Director of a local center for translation studies and research called CEIT (Centro de Estudios e Investigación sobre la Traducción). She researches Machine Translation and has presented her preliminary findings in several national and international events in the last 3 years.

For more information visit www.rwtraducciones.com.ar.

Rosana can be reached at rosana@rwtraducciones.com.ar

acquire that subtlety? Isn't it possible that the first time he notices a strange expression he will change it, the second time he will think "this sounds familiar, I saw it somewhere else," and the third time he will already assume "this is the way it is usually said"? What kind of translators will be formed in such a process? What will the threshold of quality be in the future?

Is Creativity at Stake?

As another issue, this profession, which was born as an eminently creative and expressive one, is being somehow jeopardized nowadays by all this technological progress. The possibility of being creative in our everyday tasks is becoming more and more limited: we have to follow the glossary, we have to respect the client's preferences, we have to imitate the style in the TM, we need to use Neutral Spanish (if there is such a thing), we have to unify the style of all the translators in the team... and now, we have to post-edit texts that have been automatically translated.

These new work modalities estrange us from the final result of our work. Many of us work in high-volume projects of which we only see a small part. We rarely get to know what the final destination of our work was. We just press "click" and send it, having no authorship rights over it. We are increasingly more involved in a numeric rather than a communicative process: words, hours, dollars counting.

After learning and understanding better what MT is and how it works, I have come to the conclusion that, as a translator, MT really does not worry me so much. I am much more worried by the overall automatization of the daily translator work. So finally, I'd like to share with you a paraphrase of another text by Marc Auge [1995] about technology, which I think can well be applied to the translator's profession, and explains somehow why I, as a

translator, felt the need to do research on this topic, the maximum expression of the industrialization and automatization of our work:

Only by intensifying the relationship with the technological instruments, we'll be able to control them. If we understand how they work, we'll feel less alienated by them. The new humanism is just that: forming people not as consumers, but as creators. Forming them so that they can control the instruments. Forming them to create.

I think the bottom line is: a translator's attitude should not be one of rejecting progress or opposing technology—not at all. What we all should be involved in is understanding technology, using it responsibly and productively for our benefit to the greatest extent possible, and helping clients and users become aware of its benefits and limitations. □

References

- Auge, Marc. *Fictions fin de siècle*. Fayard. Paris. 2000.
- Auge, Marc. *Introduction to an Anthropology of Supermodernity*. W. W. Norton & Co, 1995.
- Champollion, Yves. "Automated Translation: The Next Frontier." In *Translating and the Computer* 24, ASLIB, November 2002.
- Schwalbach, Ursula and Zearo, Franco. "Machine Translation: Translating Automation into New Opportunities." In *The ATA Chronicle*, Volume XXXV, Number 5, May 2006.

What we all should be involved in is understanding technology, using it responsibly and productively for our benefit to the greatest extent possible.

Call for Reviewers

Do you work as a translator and/or interpreter? Tell us about it!

Send reviews of your new, favorite, even most-hated language technology software to roomynaqvvy@gmail.com and roomynaqvvy@yahoo.com by the submission deadlines. Your colleagues will thank you!

Trados Tip

by Tuomas Kostiainen

Three issues ago I wrote about translating Excel files (Volume 1, Issue 4)—in particular, how to hide part of the source text from translation when translating in TagEditor. However, that was only a part of the whole story, and there's much more to know about translating Excel files with Trados and TagEditor. This time we are going to take a look at the SDL Trados Filter Settings application, primarily in the context of Excel files.

First, there are a few basic things to remember when translating Excel files. If your Excel file is really big (over 5 MB or more than 15 worksheets), you should split it to keep your system's response time reasonable. Also, do note that some file content is not presented for translation in TagEditor, such as some hyperlinks, text stored in external files, clip art or camera/scanner pictures, pivot tables and charts, and cell names and labels. For details on these and other limitations, see the Trados 2007 File Formats Reference Guide that came with your installation.

When you open an Excel file in TagEditor, the file is automatically converted to TradosTag format (TTX) by the SDL Trados Filter for Excel. In most cases, you don't need to worry about the filter, its settings, or any other part of the whole process. The default settings will probably work fine. However, there might be situations when the automatic conversion doesn't do the job exactly the way you would like, and for those cases, it is good to know whether you can do something about it and if so, how.

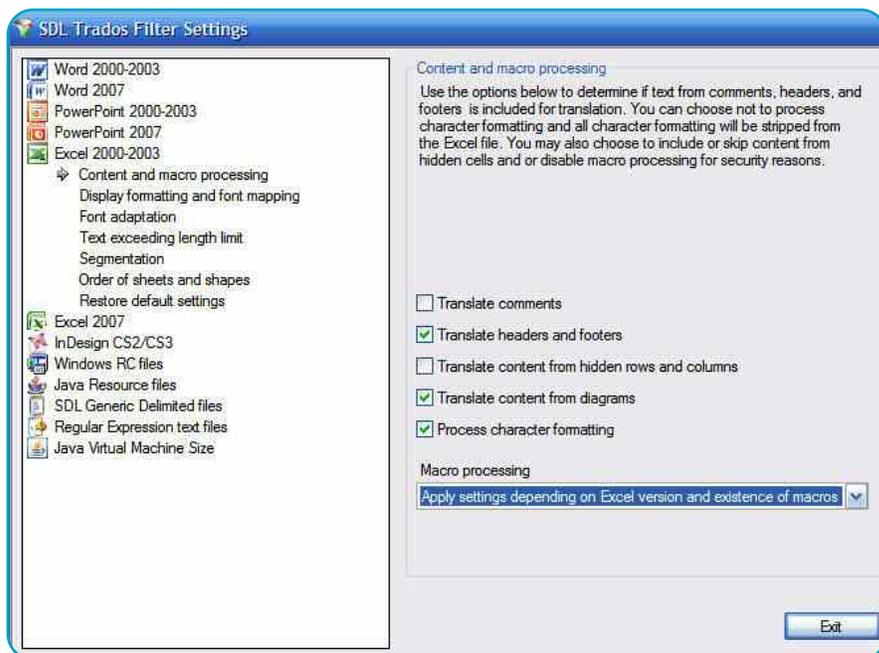
It's a very good idea to take a look at the Filter Settings application even if you don't have a burning desire to change anything right now. The first hurdle, however, is to locate it. I think most Trados users have created a desktop shortcut for their most commonly used Trados applications and have forgotten that there are some "unused" applications left behind in the tortuous SDL Trados program selection path.

Anyhow, practice your eye and hand coordination by selecting Start > Programs > SDL International > SDL Trados 2007 > Filters > Filter Settings, or access it directly from SDL Trados Synergy. The SDL Trados Filter application has filter settings, for example, for Word, PowerPoint and InDesign files, so while you are at it, also take a look at the settings for these other file formats so you will know what is available when needed.

With the Excel filter, you can change how certain types of content of Excel files are converted into the TradosTag format. The filter includes seven main options which you can select on the left side of the window (see Figure 1). These options are:

Content and macro processing – Select here whether comments, headers, footers, hidden rows/columns, and diagram content will be included for translation, and how to deal with character formatting and macros. Note that the "Translate content from hidden rows and columns" option is unselected by default which means that any hidden text in Excel file will not be presented for translation in TagEditor, as I mentioned in the Volume 1, Issue 4 article.

Display formatting and font mapping – Select whether formatting (bold, underline, etc.) is actually visible in



TagEditor or if it is only indicated by tags. This also includes font mapping options for complex script and Asian languages.

Font adaptation – You can change original small font sizes to appear bigger in TagEditor just to make the text easier to read. This does not affect the actual font size in the Excel file. There's also a similar option to make font colors more visible.

Text exceeding length limit – You might not have known this, but the Excel worksheet names may not be longer than 31 characters. If your translated worksheet name is longer, the TTX file can not be converted back to Excel format. This is frequently the culprit for cleanup problems with Excel TTX files. Here you can select how the filter should handle names that are too long. Also, remember to run the Excel verification plug-in (or SDL Trados QA Checker with applicable Excel settings) before cleaning up the TTX file because it will alert you if any of the names are too long. You can find both these plug-ins in TagEditor by selecting Tools > Plug-ins.

Segmentation – Select whether segmentation should be applied after an Excel paragraph break tag. Normally,

the filter converts these to line breaks (
 tags) which Trados Workbench, in turn, does not see as segmentation marks. Note that you can manually enter hard return segmentation breaks into the TTX file in TagEditor by placing the cursor in the appropriate location and pressing Enter. This creates a segment break but will not have any effect on the target file formatting.

Order of sheets and shapes – Determine the order in which worksheets, cells, shapes, etc. are processed for the TTX file.

Restore default settings – Last but not least, here's your panic button to restore default settings.

Also, after you have converted the translated TTX file back to Excel format, remember to check that all the content has been translated and formatted correctly. □

Tuomas Kostiainen (tuomas@jps.net) is an English to Finnish translator and Trados trainer, and has given several Trados workshops and presentations. For more Trados help information, see www.finntranslations.com/tradoshelp.

LTD Wine & Cheese Networking Event

Friday, November 7
from 5:30-6:30 p.m.

We hope that you will join us!

Please visit

www.atanet.org/conf/2008/register.htm

for information and
registration form.



Language Tech News Submission Guidelines

Submissions should be no longer than 2000 words. (If you have a piece that is longer, contact me and we can discuss options.) Send all articles and items of interest to roomynaqvvy@gmail.com and roomynaqvvy@yahoo.com in Word (.doc or .rtf) format and be sure to include a short bio at the end, ideally containing your e-mail or website contact. Please send all communications to both the e-mail addresses and mention "ATA LTD Newsletter December 2008" in the subject of the e-mail.

We also gratefully accept reprints of articles you may have published elsewhere, so be sure to offer those as well.

Next IssueVol. 2/No. 4
Submission DeadlineTuesday, Nov. 30
Planned Publication DateDec. 2008

2008 Conference Topics of Interest to LTD Members:

As of October 22, 2008

SESSION	TIME	DESCRIPTION
Seminar J	Wednesday 2:00-5:00 pm	<i>How to Get the Most Out of Your Computer: Tips for Small Business Owners</i> Carey Holzman (All levels, presented in English)
LT-1	Thursday 11:30 am - 12:30 pm	<i>Blogging: How and Why</i> Corinne L. McKay (All levels, presented in English)
LT-2	Thursday 2:00 - 3:30 pm	<i>Translation Support Tools Forum, Part I</i> Alan K. Melby (All levels, presented in English)
LT-3	Thursday 4:00 - 5:00 pm	<i>Translation Support Tools Forum, Part II</i> Alan K. Melby (All levels, presented in English)
LT-4	Friday 11:00 am - 12:00 pm	<i>Free and Open Source Software for Translators</i> Dierk Seeburg (All levels, presented in English)
LT-5	Friday 2:00 - 3:30 pm	<i>How to Keep Your PC in Tip-top Shape</i> Carey Holzman (All levels, presented in English)
LT-6	Friday 4:00 - 5:00 pm	<i>Language Technology Division Annual Meeting</i> Dierk Seeburg
LT-8	Saturday 2:00 - 3:30 pm	<i>Translation Technology's Ring of Power: One Tool to Rule Them All . . . and in the Darkness Bind Them?</i> Jost O. Zetsche (All levels, presented in English)
LT-9	Saturday 9:00 - 10:30 am	<i>First Date: A Dialogue Between Translators and Machine Translation Developers</i> Laurie M. Gerber (All levels, presented in English)

2008 Conference Topics of Interest to LTD Members:

cont.

SESSION	TIME	DESCRIPTION
LT-10	Saturday 11:00 am - 12:00 pm	<i>Simplified Technical English: Cost Savings and Quality Assurance for Translations</i> Berry Braster (Intermediate, presented in English)
LT-11	Saturday 11:00 am - 12:00 pm	<i>Automation and Alienation: The Effects of Machine Translation on Current Working Practices</i> Rosana N. Wolochwianski (All levels, presented in English)
LT-12	Saturday 9:00 - 10:30 am	<i>Best of Both Worlds: Combining Windows and Linux</i> Topi K. Junkkari (All levels, presented in English)
LT-13	Saturday 2:00 - 3:30 pm	<i>Mixing Computer-assisted Translation and Machine Translation: The Good, the Bad, and the Ugly</i> Giovana C. Boselli and Cristina Silva (Intermediate, presented in English)
LT-15	Saturday 4:00 - 5:00 pm	<i>Where is Language Technology Going?</i> Jennifer DeCamp (All levels, presented in English)
J-1	Thursday 11:30 am - 12:30 pm	<i>Optimizing Terminology Mining</i> Akiko Sasaki-Summers (All levels, presented in Japanese)
SL-2	Thursday 2:00 - 3:30 pm	<i>Software Tools for Slavists, Part I</i> Becky Blackley, Fred Grasso, Jennifer L. Guernsey, and John W. Riedl (All levels, presented in English)
SL-3	Thursday 4:00 - 5:00 pm	<i>Software Tools for Slavists, Part II</i> Becky Blackley, Fred Grasso, Jennifer L. Guernsey, and John W. Riedl (All levels, presented in English)

For more information, go to www.atanet.org/conf/2008/learn.htm