ATA gathers in Orlando

Report on the 41st Annual American Translators Association Conference, Orlando, Florida

Written by:

Carolyn
Quintero

The annual ATA conference gets the largest attendance, and the highest number of papers, of any regular translation conference series in the world. It has therefore become the major translation market place for an exchange of views and information. As always International Journal for Language and Documentation brings you a full report.

PRE-CONFERENCE WORKSHOPS. There was quite a selection of pre-conference workshops offered on Wednesday all day (six in the morning and six others in the afternoon) before the official opening of the ATA conference on Thursday morning. One problem, however, is that one of the important sessions, the only all-day workshop, did not get on the Preliminary Program. The result was that we did not know about it until after signing up and paying for other seminars. Not a flawless launch for the annual conference, but still, as conferences go, this year's event was a pretty good one.

WHICH SPANISH IS STANDARD? The erudite Raúl Avila offered one of the Wednesday morning pre-conference seminars. Avila related quite well to his audience and was obviously a scholar of great merit with an interesting project to tell us about. However, well into the first hour of the three-hour session, he still had not introduced the topic, but was busy getting to know his audience, many of whom had much to say about their own background and interests. Fascinating, but hardly what we all paid $50 to hear. Once the audience got beyond the voicing of their own opinions on terms and pronunciation, Avila's study of what is to be taken as standard for Spanish, both in regard to pronunciation and to lexical items, was well crafted and potentially quite useful. We shall look forward to the final results and to hearing from Dr Avila at future conferences.

A SOLUTION? Another three-hour Wednesday seminar was the one by Leticia Leduc, "Use of Terminological Methodology in Translation: A tremendous solution to a difficult problem." Ms Leduc began with an historical and theoretical overview of the fields of translation and terminology, identifying at length some of the differences between the two. It baffled this listener as to why she bothered when, clearly, most of the audience had paid to hear about the "difficult problem and the tremendous solution". The seminar purported to demonstrate that the use of terminological methodology in translation solves the problem of how to emulate the specialist in the field in which we are translating. The final half of the session was spent on a series of exercises based with the audience divided into three teams. Each team was loosely instructed to use different translation aids: one using a simple dictionary, one using only an unorganised set of short texts in which they were to perhaps find some of the needed vocabulary. It was not clear what aids the other group was using, perhaps none. Some in the confused audience had come that afternoon perhaps expecting to hear about a simplified computer-based program to manage the terminology we have almost all collected willy-nilly. In fact, neither the point of the team exercise nor the instructions for the groups were very clear. A great deal of time and effort was expended in each group to produce results that were then read out loud to all, with the presenter critiquing. The problem was that the persons reading were not identified as to which "tools" group they represented, so that no point was effectively made. Many vocabulary items that stump translators were mentioned and rapidly clarified, and are presumably found somewhere in the plethora of materials handed out. Ms Leduc has extremely valuable expertise as a consummate professional, both as a scholar and a practitioner. Perhaps in a more organised effort she will manage to communicate more of her knowledge to us in a future seminar. I, for one, would sign up again, but with the hope of a) less theory , b) more clarity and c) less time spent illustrating how much we do not know and what the wrong ways of going about things are.

Brief glimpses of the seminar by Xosé Castro showed a large audience extremely engaged by the presenter who, with many humorous asides, was cajoling translators to make fuller use of the tools they have at hand.

OFFICIALLY BEGUN. Thursday, bright and early, found our president Ann Macfarlane conducting the opening session with her characteristic eloquence, authority and humour. Linda Sivesind, NFF (Norway), the cheery vice president of the international federation of translators (FIT), brought greetings to us who are "by nature bi-cultural or multi-cultural and by nature invisible, if we have done a good job." Sivesind made the point that we need to make ourselves more visible, and reminded us that September 30 is International Translators Day, the birthday of St Jerome, the patron saint of translators, and the beginning of International Translation Week.

OFFICIAL KUDOS. Courtney Searls-Ridge headed up a committee that read 44 books in nine languages in order to choose a winner for one of the prizes (see the ATA Chronicle for a full list of recipients). Gabe Bokor, a highly esteemed ATA member, received the Alexander Gode medal; Peter Meinic was honored for a new, "readable and quite beautiful" translation of an ancient Greek play; Jessica Cohen, a Bloomington, Indiana student, won the student prize.

HIGH TECH PIONEER. Bokor has been in ATA since the late seventies and was a word processing pioneer, demonstrating electronic data communication and tools to ATA members in the 1980s. His acceptance speech gave an historic overview of technology development with respect to translating. Bokor exhorts us to fine-tune our marketing strategies, as marketing to all and sundry too often "resembles an oriental bazaar." He cautions that machine translation "does not do the job" but admonishes us that we cannot and should not ignore progress: "technology is here to stay."

MATURE, VIGOROUS AND POWERFUL. Ms Macfarlane reviewed ATA progress over the past year, referring to us as a mature, vigorous and powerful group. We have agreed, it seems, to disagree. A case in point is the question of international setting for examinations. We have numerous Chapters, Regional Associations and 12 Divisions. We are able to let something we love come to an end (the Sci-Tech division), which is, apparently, a sign of real maturity. We honour our past, and we invest in our future (testing program, mentoring, distance learning). We offer benefits to our members, such as unique marketing services in our directories and a board initiative to do targeted marketing. We know how to enjoy ourselves (witness our poetry and prose readings at the conference and our parties!). Macfarlane comforts us with the assurance that board members do not just sit up on the stage like flower pots.

NEW BLOOD. Tom West, our new president, spoke briefly, requesting feedback above all. We shall all look forward to the talented Mr West's presidency of the ATA, from this conference on.

TECHONOLOGY SESSIONS. One of the most interesting parts of the annual ATA conference is the daily presentations in the Translation and Computers (TAC) sessions. Especially gratifying this year was seeing that the speakers have come down to earth. No more exaggerated claims of miracles, a comfortable and reasoned analysis of where the industry stands and who can benefit from the technology under what circumstances. On Friday afternoon Alan Melby opened a session on Translation Memory Tools - Fact or Fiction? Brian Chandler (Star Transit) spoke, followed by Christina Spies (Trados) in How Can TRADOS Tools Benefit Your Localisation Projects?, then

Brian Briggs (Language Partners International) in Beyond Translation Memory: The Latest Second-Generation CAT Tools. Closing the afternoon sessions was Denise Baldwin (SDLX), speaking on (what else?) SDLX.

STAR AND MYTHS. Brian Chandler has been working with translation memory (TM) tools for more than eight years and gave us an overview of TM, contrasting database versus file pair, which is their approach in STAR. The major claim is that TM cuts down on time, but this is only true depending on several factors. Imported text, e.g. a Quark file, must be filtered to leave only translatable material, then the text must be broken up into pieces which constitutes creation of the file pair. Subsequently text is pre-translated if there is a 100% match from previously stored materials. In the final step, the translated materials must be imported back into the original file type, such as a Quark file. Good project management is extremely important. Chandler exposed certain myths or fictions. One is that "TM is easy to use". In fact, one to two weeks are necessary to learn it to a level where one is confident with 40 to 50% of its capabilities, and more time is better. A survey has shown that only 50% of the people who have purchased TM use it consistently with the major complaint being that it is too hard to use. Another myth is that TM "will automatically translate a document." Among items for which TM is "not worth it" are: correspondence, literature, advertising, small marketing pieces. Another TM fiction is that TM tools never make mistakes. Translation is an art, says Chandler, and no tool is perfect. A related fiction is that TM tools eliminate the need for review. Several things can go wrong, including formatting, faulty algorithms, currency translation problems, etc. and review is absolutely necessary. Alignment of the text could have been off, for one thing, and for another, the flow of the text could be faulty as material is brought in from several sources. Another myth is that TM eliminates the need for desk top publishing efforts. Expansion problems can arise, especially in figures, graphics, tables.

FIRST UNTRUTH. An important fiction to decry is the one saying that there will be immediate savings in time and money. In fact, Chandler says, there is a loss of money in the first few translations, with the learning curve and project setup time - and at this point many translators abandon the effort. One of the scariest fictions for some translators still unfamiliar with technology is that TM tools eliminate translators. Simply not true.

OTHER MYTHS. Other problems mentioned by Chandler in this very even-handed and helpful seminar are: Formatting is invisible during the process. Tags must be protected. File pairs must not get misaligned. Currency and date problems arise. Segmen-

tation of text varies as to language, kinds of abbreviations used, and several other factors. Clients may request a different segmentation; these may occasion paragraph and matching problems. Clients may pass along updates outside STAR. There can be low text repetition, or variations in sentences may be interfering with leveraging the previously translated materials. And, finally, there may be ownership questions.

97% APPLICATION. Despite all the potential problems, Chandler's group use STAR Transit in 97% of their work, finding that it reduces time in the long run. TM uses 70% of the time used by human translation of the same job, saving 30% time. The consistency among translators and among translations is enhanced, and terminology databases should be requested from the client because they are needed for the TM tool. Chandler's TM tool actually checks for format at the end of a project. Most importantly, there are cost savings overall.

MONEY TALKS. If a $22,000 translation project can reap a 20% match from stored memory, then the project will cost $19,500. If 40% of the text blocks come up matching, then the cost will be $17,500. Even with all the pitfalls mentioned, these tools are needed to compete in consistency and cost and they also help significantly with project management. Updates are easier using TM. As to TMX format, it is "almost there" says Chandler. TM is not a miracle, and one needs to invest time in it.

TESTIMONY. A user of STAR transit gave her report. Muriel Wang says that the full version, which includes project management, takes one week to learn and that the light version takes about three days.

TRADOS. Christina Spies from TRADOS in Alexandria, Virginia, says her company has a 70% market share, and there is a bundle of products integrated with Translators' Workbench. She reviewed the steps: 1. populate memory (optional), 2. batch process (i.e. compare new document to memory and extract portions) with a mere click of the mouse, 3. send off untranslated portions to vendors (freelancers). Spies reminds us that TM is not just a database, that it has a "fuzzy engine", but joins Chandler in agreeing that these tools are not perfect. They are, however, fairly user friendly in that they integrate with MS Word and integrate into a terminology system, for those of us who can be said to use one. Her product supports TMX, that is, it can get the contents out of memory and use this material with other software. Best application occurs, says Spies, where a) consistency is important, such as the medical field, b) several translators work on one document, c) there are updates to be dealt with, d) there are frequent releases, such as a newsletter, even if it's only ten per cent recycling, e) recycling happens across document types, such as HTML web pages converted into a printed manual or help files in rtf format are turned into a hard copy manual, f) across projects e.g. a Word manual may be converted into an Excel manual.

DEJAVU'S FREE UPGRADE. Then DejaVu came on the scene in the person of Brian Briggs, who also advocates this recycling of databases. His product handles TMX and TRADOS interface, and has the unique feature, he says, of building a lexicon. DejaVu looks for terminology to build a lexicon, and if not found in the new terminology database, it goes to its own terminology database. DV gives us the number of occurrences of a word, as part of the decision of what should be included in the terminology database. This also is a tool that can assemble translations from portions of text, not just the exact and fuzzy matching of pairs of sentences. In fact, this feature is included in the colour coding of DejaVu:

green is exact match; magenta is fuzzy match; blue is "assembled" (where only a part of a sentence matches). Briggs spoke of the next generation, which will integrate more with the client, improve usability with additional linguistic technology, distributed database capability for translators working on one project to have access via the Internet. The newer version 4.0 will have multiple file formats; revision history; more of the feel of MS Office, and will be available free of charge to those purchasing version 3.12 at this year's conference and owners who purchased at last year's conference, says Briggs. The new version will be available at the end of the year.

SDLX. The fourth speaker of the afternoon was Denise Baldwin, on SDLX. The SDLX TM tool makes translation memories portable between products via the industry standard TMX (Translation Memory Exchange), she says, and will continue to support any relevant standards as they become available, including TBX (Terminology Base Exchange). SDLX is compatible with TRADOS and IBM Translator products. All four speakers were quite interesting and gave us one of the most relevant and important afternoons.

SOFTWARE VENDORS TAKE HEED! An important point for the vendors of TM software is that we ATA users are key in promoting their product. It would certainly be in their best interests to offer hands-on training or troubleshooting sessions at the conference for those of us who have purchased their software and may not be using it fully because of problems we have had with it. We are your best marketing tools! A word from one of us to another translator thinking of investing in TM will often be decisive in the decision he or she is making. So it would be smart of you not to charge us hundreds of dollars to learn to use the product we have already purchased! It is wrong thinking to add this punitive cost when we have paid for the product and you could so easily offer us a helping hand at no charge where you find scores of us already gathered in one easy location.

At 8:30 am on Saturday, Jackie Murgida of Lernout & Hauspie spoke on Machine Translation: Getting Down to Business. "Hype aside, machine translation is settling into its proper place," said the abstract for this talk, "not as a replacement for human translation, but as an important part of the profession as it is practiced today." Dr Murgida conducted an effective illustration by showing the audience a page of raw output of Arabic text which left most of us clueless as to type of document, content, etc. After some guessing by the audience, we got to see the same document passed through machine translation and quickly were able to grasp its nature (an interview) and most of its content. Murgida was showing us the usefulness of MT as a "sifter" or "triage" tool. Its best use is for "gisting", for taking large amounts of material with unknown content, where we need an idea of what we're dealing with. Although not yet a product, this device will allow us in the future to use translation memory (TM) where possible, then apply

machine translation (MT) to the remainder, using dictionary and lookup components. Post-editing, according to Murgida, is still entirely too cumbersome. "Time is one thing but aggravation is another!" she says, not being someone who wants to post-edit MT. "And they want to pay you less, not more!"

UP & DOWN OR BACK & FORTH? Murgida advocates a left and right lineup of source and target texts, saying that up-and-down texts are hard to use. Searching for names in foreign articles will be a helpful part of the useful browsing, indexing and searching tools for choosing material to be translated. Nokia uses MT for Finnish for the triage and browsing functions, she reports. This session was not really about the pros and cons of MT, but rather a report on how it is being used to enhance productivity and advance the profession at large.

FOR EVERYONE. In this writer's opinion, all ATA members should attend these talks and get informed ASAP about the technology in our field, whether we are users or not. Speakers like Murgida can bring the message to us in a way everyone can understand and relate to.

READY TO REPLACE HUMANS? Stefen Lampert of XTRA International Services, a hesitant but effective speaker, will tell companies who want to implement MT solutions how to do so. After presenting his resume on a slide, Lampert told us that he has done an in-house comparison of TM vs MT usage, and that his company is not using MT at present for client projects. He mentioned that we are at the very beginning of use of voice recognition and MT systems for one-on-one communication between e.g. a Japanese and a US businessman (presumably Lampert includes women, too). He spoke of the two levels of technology, the lower level being the use of dictionaries with little analysis of sentence structure (which is "Direct") versus the "Transfer" level which involves interpreting meaning with word and sentence structure analysis to produce an image (metalanguage) in between, then proceeding on to the target language.

THE PLAYERS. There are three kinds of players in the market: 1) MT developed for and by large corporations, 2) low cost MT such as Altavista which is typically offered free, 3) commercially available MT engines with subject related dictionaries.

WHAT PROGRESS HAS BEEN MADE? What's the progress in the technology? asks Lampert, and answers by quoting an authority at the conference: "Virtually none. The major players have left the playground". Lampert concedes that many of the major players have lost interest, but says that machine processing does work much better now. Computers have improved greatly with both increased knowledge and processing power, so the time involved is less, and this is a key factor. Although it may seem there has been no development in the past ten years, this processing power has multiplied and the focus has now changed to having the right setting and specific framework required to reap the benefits of the technology - an important insight on the field.

WHICH LANGUAGES? Most MT vendors support FIGS (major European languages), plus Japanese and/or Portuguese, in file formats including RTF, HTML, SGML, PageMaker, FrameMaker.

WHAT ARE THE ISSUES? Lampert identified several preliminary issues: Does the amount of the documentation justify MT? Not even a few thousand pages does. What is the type of documentation? Styles must be similar, e.g. a group of maintenance manuals. What are the quality requirements? Is the translation available free of charge? A high cost should give a better final product. Is an "MT-like" style acceptable for the particular project? Would a mistranslation be risky? This would be the case for a legal document, says Lampert. It's important to define the appropriate quality (a concept new to many translators). Other factors are: how the source materials impact the MT output, that is, are there typos and other characteristics that will cause problems for MT. The solution to this is Controlled Language. A comprehensive terminology database is key, but is a different kind of database also necessary? Can we extract and define terminology before translating?

QUALITY. Lampert cautions that post-editing is part of quality assurance, not an option, and says that post-editing requires higher qualifications than translation itself. We need to keep post-editors from getting frustrated, motivating them to improve the system with their input.