

## Translators Still in The Loop

**An MT workshop and a translator's conference held back-to-back in San Diego set in clear relief the sometimes uneasy symbiosis of translators and machine translation.**

*San Diego, California* – The International Association for Machine Translation (IAMT) scheduled its first MT Evaluation Workshop and MT Showcase to coincide with the 33rd annual conference of the American Translators Association (ATA) here at the lush tropical enclave of the Mission Bay Princess resort. In retrospect, it seemed like an obvious decision – the two disciplines remain so inextricably intertwined.

The burgeoning IAMT, scarcely eighteen months old and already counting five-hundred plus members, is the umbrella organization for the Association for Machine Translation in the Americas (AMTA), the European Association for Machine translation (EAMT), and the Asian-Pacific Association for Machine Translation (APAMT). The ATA represents four-thousand professional translators in North America.

With the IAMT Workshop and concurrent Showcase taking Monday and Tuesday for its reckoning and the ATA laying claim to Thursday, Friday, and Saturday, both camps got plenty of chance to inspect each other, although of course not all Workshop participants stayed on and not all translators came early. For people who stayed the entire week, two conclusions were obvious: one, translators are still very much part of the process of machine translation, and two, MT is irrevocably becoming a fundamental part of the technical translation process. A wager: there will not be a single technical translator in ten year's time who does not speak of "my MT system."

The IAMT Evaluation Workshop sessions attracted an eclectic crowd that could be roughly broken down into researchers, vendors, and users, although you could argue that the latter group includes two factions who do not necessarily share the same interests: what you might call "super-users," documentation managers and others who are turning to or considering MT for commercial reasons but who are not necessarily linguists, and "end-users," the translators who are hired to post-edit. This distinction can be made because not many translators use MT for their personal benefit – yet.

In any case, the Workshop sessions saw plenty of lively interchanges, ranging from esoteric theoretical debates ("blackbox" versus "glassbox" evaluation theory) to pragmatic matters, such as WordPerfect compatibility (high on everyone's list) and the importance of macros for post-editing.

### **A profusion of systems**

AMTA president and Workshop coordinator Muriel Vasconcellos noted in her opening remarks that the five-hundred members of the combined MT associations had a substantial chance of charting the direction of MT in the future. "Serious, mature systems are now running on PCs," she continued, "meaning we can finally say goodbye to the mainframe. MT is coming to our desktops. The question is: What do people need to know about MT to understand it? There are now thirteen English-to-Spanish systems out there. How do you choose which one to use?" Following Vasconcellos, Makoto Nagao reported that ten commercial MT systems are currently available in Japan; ten more are under development. "We are entering an interesting period," was his succinct observation.

As the discussion on evaluation unfolded, you might have been forgiven for thinking that evaluating an MT system is almost as complex a matter as developing one. New Mexico State University's Yorick Wilks called for purpose-oriented evaluation: What will the translation be used for? Publication? Content-scanning? He dismissed the notion of

“glass boxes;” the inner workings of a system should not influence how you evaluate it. While Margaret King of the ISSCO (Geneva) supported the notion of purpose-oriented evaluation, she pointed out that you cannot judge the system simply on the basis of the output: “Was the text mis-translated? Or was the source ambiguous? These are two separate problems.” She maintained that the criteria for evaluating machine and human translators were “radically different;” this by definition took into consideration the “methodology” involved. King was seconded in this respect by metal user Doris Albisser of the Union Bank of Switzerland.

Wilks refused to let this pass, maintaining that purpose-oriented evaluation and separate criteria for man and machine were two conflicting opinions. The CEC’s Loll Rolling took the diplomatic position that evaluation should be based on requirements of the user community, while John Benoit of the Mitre Corp, speaking from the point of view of an independent tester, said he based his evaluations on the required end result. For him, that could be relevance scanning, content understanding, or publishing.

Muriel Vasconcellos called upon the members of the MT associations to work together to develop a standard suite of two-thousand test sentences for evaluating MT systems. Alluding to the fact that DARPA keeps its test corpora secret from DARPA-funded developers, she said such sentences should be freely available. “Sure, you can pad the dictionaries with the vocabulary you need – but syntax you can’t fake. Either a system can do it or it can’t.”

### **But is it compatible?**

If there was any one common concern among the users and potential users present, it was compatibility: compatibility with existing wordprocessors and publishing systems and compatibility with translators. Except for MT systems used purely for content scanning, today’s systems require some form of human intervention – and that human is usually a translator. This implies systems must reflect an understanding of how translators work. To use the distinction once made by Kent State University’s Gregory Shreve, effective systems will need to be translator-centered, not translation-centered.

What translators appear to dislike most about post-editing is making repetitive mechanical changes. Vasconcellos suggested one way of cultivating translator compatibility was to provide wordprocessor macros for automating common post-editing tasks, such as switching words, changing adjectives to adverbs, deleting definite articles, and making such changes as “in order to study” to “for studying” and “of the group” to “the group’s.” These simple timesavers can take much of the drudgery out of post-editing.

Thinking of the future, CompuServe’s Cole Harrison raised the issue of forward compatibility. If he spent five years building the dictionaries of one system, could he be assured that he would be able to move that lexical data onto another system? He was given no reassurance of this, although one member of the audience lightheartedly suggested making it a licensing condition and throwing that responsibility in the lap of the vendor.

### **Numbers speak volumes**

For the ATA conference participants, varying from those freelance translators who could afford the US\$120 a night for a double at the swank San Diego Princess to translation company representatives and corporate translation managers, machine translation is clearly no longer on the fringe. There was a full house on Wednesday for a day-long seminar on machine translation for translators led by Joann Ryan and Veronica Lawson while the MT-oriented sessions the rest of the week were very well-attended. But translators are not starting to take MT seriously because of some dazzling parse they have seen in vendor demonstrations in the adjacent Sunset room; if anything, it is numbers. Robert Hensley of the AT&T Document Development Organization, for example, who participated in the MT in Practical Applications panel, explained that AT&T would be translating an eyebrow-raising 350,000 pages of technical documentation this year with MT. AT&T uses MT to translate to and from English, Spanish, German, Russian, French, Portuguese, and Japanese, using systems from Logos, Intergraph, Sharp, and Tovna. Hensley explained that not only did MT significantly increase translation capacity and lower costs by more than half, it resulted in near one-hundred percent consistency in

terminology, something manual translation could nowhere near approach. AT&T hires free-lance post-editors, who either work on-site or send and retrieve material via modem.

To translators, is MT then a threat or an opportunity? Insofar as the fear is unknown, they might be forgiven being apprehensive; several, however, rose to the occasion and asked pertinent questions, such as: Who would train them to post-edit? and How much did companies pay for post-editing? The Practical Applications panelists, all big MT users, did agree on one thing: the best if not the only suitable post-editor is a well-qualified translator. Post-editing, at least in its current form, requires an understanding of translation and familiarity with both the source and target texts. Where translators go, agencies are sure to follow: there are now translation companies which specialize in post-editing.

### **An MT evangelist**

Relaxing after long and hectic days at the helm of the IAMT Workshop sessions, Vasconcellos commented that the MT Associations are helping break down the traditional barriers between academic researchers, commercial developers, and users. Commercial developers have not always understood the problems faced by users; users in turn have not understood the formidable challenges faced by developers; and researchers, who use MT systems to test linguistic theories, have tended to ignore the former while regarding the systems offered by the latter as primitive.

“Heuristics,” Vasconcellos maintains, “has been a dirty word in academic circles. It’s good for the researchers to have contact with suppliers of operational systems.” Forums like the Evaluation Workshop are the only place that people like Systran’s Denis Gachot and IBM’s Peter Brown are likely to meet. Systran’s linguists on a whole do not publish in academic journals and Peter Brown, well, he has his DARPA evaluations to keep him occupied. An occasion like this is, figuratively speaking, the only opportunity Brown might have to ask Gachot whether Systran does in fact parse.

For Vasconcellos, the establishment of the MT associations and the holding of the MT Evaluation Workshop and Showcase are the consummation of many long years of promoting MT, trying to get it “out of the closet.” The career of Vasconcellos, who has a PhD in theoretical linguistics, is a unique one, encompassing her participation in early MT research at Georgetown University in the 1950s followed by her long tenure at the Pan American Health Organization (PAHO), where, as chief of translation and terminology, she has both supervised a team of professional translators and overseen the in-house development and implementation of PAHO’s translation systems, SPANAM and ENGSPAN. Clearly devoted to MT, she remains charmingly paradoxical about it: pragmatic yet at the same time idealistic. Musing at the heavy funding going towards unproven techniques such as interlingua-based and knowledge-base MT, she wonders whether investments in mature systems like Systran and Logos might not be more sensible in the long run. Yet she also exhorts researchers to spend more time analyzing discourse markers and “extrasentential” information. In any case, her work at PAHO having reached a logical juncture, Vasconcellos is taking an early retirement from the PAHO this November to be able to devote more time to the MT associations and one or two writing projects.

### **The new alchemists?**

The final word of the day must go to University of Southern California Information Science Institute’s Eduard Hovy, who, in the final Workshop panel, cautioned us not to let MT research become the modern-day alchemy. He recounted with a twinkle in his eye how one day, while waiting for PANGLOSS (his DARPA-funded MT project) to parse a sentence, he fell asleep and met the alchemist Paracelsus in a dream. The famous man inquired “Had people finally succeeded in making gold out of base metals?” “No,” Hovy replied, “not in the least, but we now have a science called Inorganic Chemistry and another called Organic Chemistry and they both do all kinds of wonderful things, from creating plastics to fabricating dyes and making medicine.” On waking up, Hovy then started wondering: what would someone 200 years into the future answer if he were to ask: Had people finally succeeded in building an MT system? Probably a similar answer: No, we don’t have a universal MT system. But we have our email reader, our court proceeding processor, our portable travel translator, newspaper summarizer, etc. “It could be a mistake to be obsessed

with one goal,” concluded Hovy. “Maybe we should be thinking in terms of a plethora of systems. If we concentrate just on the ‘gold’, that is, the perfect MT system, we might never find it!”

COPYRIGHT © 1992 BY LANGUAGE INDUSTRY MONITOR