PHOTOGRAPHY: LOUIS ROSSETTO

MAGHI KING
ON MACHINE TRANSLATION

For the past thirteen years, Margaret (known to all in MT land as Maghi) King has headed Geneva's Institute for Semantic and Cognitive Studies (ISSCO), an independent research establishment founded by Italian businessman-philanthropist Dalle Molle and dedicated to furthering the cause of "human ecology." She also chairs the Association for Computational Linguistics' European chapter and is vicechair of the European Coordinating Committee for Artificial Intelligence. As director of ISSCO, she was responsible for the coordination of Eurotra, the European Community's mammoth nine-language machine translation project, from its launch in 1978 till early 1983.

Born in Welsh-speaking Gwynedd and raised in Manchester, England, Maghi found out young that a foreign language was not just a translation of her own mother tongue but the key to a foreign world picture. After studying classical languages and philosophy at Oxford, her first job was teaching secondary school Latin. Soon weary of the low pay and status Britain's teaching profession commands, and with a young family to support, she quit to train as a systems programmer with UK chemicals giant ICI. In 1969, she became a researcher at UMIST's Regional Computing Center, where her interest in machine translation was born and further nurtured through her correspondence with MT pioneer Yorick Wilks. In 1974, Maghi came to ISSCO, whose range of activities includes theoretical and computational linguistics, natural language processing and artificial intelligence for machine translation.

What of the current state of MT?

"The products already on the market – they're all very far from perfect, yet they sell because the need won't go away. Nor can you single one out as better than the rest. The commercial products – Systran, ALPS, Weidner – all vary. They've been built for different purposes. Systran is aimed at general purpose batch output — large volumes of all sorts of text in a hurry. Which is almost bound to mean a lot of postediting just for accuracy, let alone style. Weidner can do the same type of batch processing but has to be tuned to a particular type of text. ALPS is interactive – and therefore slower – but you can control what's happening if you want total accuracy first time round. The only imminent competition I can envisage in the commercial sector is from Siemens – their Metal program. Which isn't on the market yet.

"Of other systems still being worked on, Rosetta looks to have a nice clean methodological basis. But it's not meant as a competitive product. DLT – with the Esperanto interlingua – is one I'm very skeptical about. I'd be surprised if it's capable of large-scale translation – semantics and world knowledge techniques make for notoriously unstable systems. As for Eurotra, now that's a completely different animal – it's bigger and more ambitious than anything else around. There's simply nothing else designed to do batch translation of all kinds of text in 72 language pairs. But I hesitate to predict what will come out of it."

As a multilingual project, why wasn't Eurotra developed as an interlingual system like DLT?

"None of the people working on the initial specifications believed that an interlingual approach could work for a system of its size. And I agree. I'd still use transfer architecture today. Attempting to construct an interlingua for several languages is a practically impos-

sible task. It amounts to constructing a symbolic model of the world. It rests on the false assumption that all languages perceive the world and draw conclusions from it in more or less the same way. In reality, each language projects its own structure onto the world and divides it up in its own peculiar way."

Let's talk about Japan.

"They're spending money. It's absolutely great. There's the big national Mu project, now following up its extensive predevelopment stage for Japanese and English. And a new project on Asian languages, including Chinese. And the telephone interpreting project that started last year, planned for 15 years. On top of these they've got about 20 research projects in universities and industry, and about 12 commercial programs. The Japanese have taken the lead in MT – because they're not scared of investing in long-term research."

What's going on now at ISSCO?

"Right now we're investigating the use of a technique we call Constraint Propagation for MT and other types of natural language applications. Roughly speaking, it analyzes text by first extracting only the information contained in the text itself. It then takes this "first representation" and adds information based on what is known of the text's real world domain."

How does the future look for MT?

"Machine translation's here to stay because the demand won't go away. And it's growing – for both political and economic reasons. Nation states and cultural groups will continue to demand that all forms of documentation be translated into their own languages. And with the growth of scientific knowledge and economic development – especially in the Third World – the volume of documentation is bound to grow too. It's the volume that human translators just can't cope with on their own.

"The Japanese have realized this, and the West is gradually coming to see it too. There'll be no more going back to the lean years after the ALPAC report in 1966 – which knocked MT for six and halted funding overnight. For ten years afterwards there was practically no research in the West. Ironically, as a glaring contradiction to ALPAC's findings, the Georgetown Russian-English program – which ALPAC tore apart – was used throughout that period by ISPRA to translate natural science documentation. Because they needed it.

"Ever since Eurotra got started in 1978, there's been a continuous expansion of MT research, both academic and within the large corporations. And as a rule this research trickles down slowly to the commercial systems, which by definition are always based on out-of-date technology. The big problems remain: those of semantics, world knowledge – and discourse aspects arising from implicit assumptions in texts that are obvious to humans but not to machines. But sheer practical need means that MT is bound to get better."

What makes working in MT fun?

"For me, MT offers a combination of philosophy, languages and writing formal systems which I love. It has the additional bonus of letting me know immediately whether I've got the answers right or wrong."

– Geoff Pogson