MT discussed at Basle high-tech conference

Molden describes IBM experience

Computer-aided translation was one of the major topics at an International High-Tech Forum held in Basle, Switzerland, at the end of November. The forum brought together leading executives and researchers from multinational companies from all parts of the free world.

Marc Matoza, managing director of the European division of Automated Language Processing Systems (ALPS), described why productivity breakthroughs in translation had been slow to develop, but what could be done in the future to "turn the potential language barrier into an opportunity".

"Europe", he said, "is identified with a multilingual culture. This multilingual culture is not only a "way of life", it poses an important challenge to the European business community. To effectively compete against technology driven countries, such as Japan and the United States, Europe has to be able to pool its technology and marketing resources. Translation is an answer to this challenge, but unlike the accounting or inventory challenge, it has not benefited from modern productivity tools, such as the computer."

Mr Matoza said that material to be translated varied from general correspondence to highly technical and repetitive servicing manuals. "Just as the application varies, so do the technical problems (volume and terminology content) and, naturally, potential tool/solution".

Translation was a text-driven process. Modern advances, such as computer-driven word processors, were providing productivity gains in this area, and setting the stage for a new generation of word/text processing tools.

The two linguistic challenges were terminology and the resolution of ambiguity. The front of a car was called a *hood* in the United States, and a *bonnet* in England; both terms were correct in their proper environment, and the translator had to have control of this element to ensure the quality of the translation. An example of ambiguity was "IBM is a small computer company". Mr Matoza went on to describe how ALPS had tackled the problems of computer-aided translation.

Walt Molden, head of European Language Services for IBM, based at Birkerød, Denmark, described IBM's experience with computer-aided translation.

Translation of information, he said, was becoming an increasingly important activity for most multinational companies.

"If we were to look back just five or six years, we would find that very little translation was done. At that time one could find most data processing systems in what was referred to as the "glass house". In those days, an unwritten prerequisite was the ability to speak "computerese" which was a unique language sprinkled with acronyms and devised terms which bore no resemblance to any natural or national language. It was a unique language understood by the professionals in data processing and

required little or no translation. Most applications developed by the professionals in the "glass house" were created in the language of the end user so again, translation was not required.

All of this suddenly changed, however, as information processing companies like IBM began to develop and market products which were to be used directly by the end user. The user profile suddenly changed and "computerese" was no longer a viable means of communication. Suddenly we began to think in terms of the end user's "natural" language and translation became a very real requirement."

After dealing with the type of information translated, Mr Molden gave three principal reasons why some form of computer assisted translation was important: time, productivity and quality.

"Time", he said, "is the major enemy of the translator. The amount of time between the availability of the US English source and the availability of the product to be shipped is diminishing. Therefore, the time for translating is also diminishing. If the translation is not available on the same schedule as the product you are potentially jeopardising your revenue. Not an appealing alternative. So any tool which can effectively speed up the translation process is most welcome.

"As the development and innovation in new end-user products continues and grows, so does the translation volume and workload. In IBM, the amount of information translated has become increased by a very, very

significant factor over the last four years. Simply stated, we need to find ways to increase the productivity of the individual translator. Manual translation is extraordinarily labour intensive.

"No matter what pressures are exerted by time and productivity, we can never sacrifice the quality of the final product. In fact, I believe that automated translation will improve quality. This may appear as a gross contradiction since we all accept that no translation system today produces perfect results."

Mr Molden described IBM's experience, starting in the middle of 1983, in selecting, testing, evaluating and implementing a computer-assisted translation system. He described the difference between batch systems, which take a source text and process it in a batch operation without the involvement of a translator, the system-translated target language being then presented to the translator who post-edits it to produce the final translation; and interactive systems, which are structured to produce a target language text with the direct participation of a translator, tackling source language ambiguities as they arise.

IBM had elected to test an interactive system, and chose English to Italian for a nine-month test. The test consisted of an installation and training phase, dictionary or glossary building, a productivity test and a quality test.

Building and maintaining the dictionaries is a very important part of any translation system. Most systems provide a "starter" dictionary which contains common parts of speech and general vocabulary. To this, must be added user specific terminology. In IBM's case, that had meant creating their own glossaries of terms specific to the information processing industry. They started with a base of about 10,000 Italian terms and when the test concluded they had had a base of nearly 25,000 terms.

"With the dictionaries completed," Mr Molden continued, "we began our productivity test. During this phase we concentrated on numbers of words translated per hour without regard for final edited quality. We were pleased with the productivity, but not with the quality. We then conducted a second test where the focus was on the final quality level. The result of this second test gave us

a minimum 50 per cent productivity increase and a final translation which was well within our quality objectives.

Detailing some of the findings of their tests Mr Molden said it had been shown that "the first and most fundamental finding, is that it is extremely important to establish very early in the process a realistic level of expectation. You need to understand and accept the limitations of the system. If you expect the system to produce perfect translations you will be disappointed and will reject the system. If, however, you accept the system as a tool to aid the translator in performing perfect translations then you can take advantage of the benefits the system can provide.

He added that dictionary and terminology control were a key aspect of the system. "The day you employ a second translator you have created a terminology problem", he said. "Translation is not an exact science. If you give an identical source to two translators you will most likely not receive two identical target translations. Most translating organisations provide glossaries for the translators. However, the application of the glossary to the translation is left up to the translator. The translator may or may not consult the glossary. With the system, the glossary is automatically and consistently applied. Also, the glossaries can be updated and maintained by project and automatically."

Quality of the source input had also been found to be very important. Source language ambiguity was difficult for a translation system to deal with. "If the source can be simplified and controlled then the systems have a much better chance to produce a perfect translation. We hope that by using a translation system we can begin to develop better standards for the source language," he commented.

Consistency of translation was also essential particularly when there was a need to assign a project to several translators. "If all the translations are using a common system with a common terminology base you are much more likely to achieve a common translation. Lack of consistency is one of the major contributors to poor quality."

Mr Molden concluded by stating IBM's belief that the test had proved that the system can be improved and enhanced.

"This", he said, "is important for several reasons. As I mentioned earlier, we are implementing an automated system from the creation of the source to the production of the final translated product. Therefore any translation tool must be capable of being integrated into the total system.

"Secondly, we want a translation system which can continually be refined and enhanced. During the test we suggested several improvements which were implemented. These enhancements improved both the quality and the usability of|the|system. We feel that the system can continue to be enhanced within its present design architecture."

As reported in the November issue of *Language Monthly*, IBM have entered into an agreement with ALPS to develop a tailored system for IBM use. Mr Molden told the audience that they planned to begin installation of the system they had defined early in 1986 and would then continue to enhance and improve it for IBM's specific use.