

Low-cost Automatic Translation — An American Dream?

Despite enormous scepticism from academic MT research teams, a number of maverick believers are endeavouring to follow computer technology downsizing trends down the road to low-cost PC-based translation systems. So far, they are all American products. Some of them, like Texas-based Linguistic Products with its PC-Translator package, claim to be making money out of a careful adaptation of PC-power to specific translation requirements. The practical design approach is founded not on the principle that you have to model language knowledge in a computer program before getting results from a language transfer engine, but on the observation that the speed, logic and accessibility of microprocessor power can be usefully applied to that subset of language and document handling tasks which can be easily rationalized. Linguistic Products' clients, for example, tend to be the documentation wings of large high-technology companies with repetitive in-house translation loads, and not, as you might expect, the 'small' companies that might be attracted by programs sold for a few thousand dollars only.

Mainstream language industry pundits denounce this PC-based translationware as being 90% marketing chutzpa and 10% delivery—ie, they are not strictly-speaking 'language engineering' products at all. Yet the key point is that they are user-driven, and if there are in fact tasks that these systems help carry out, then it is up to users to endorse --or denounce—the cost-effectiveness, user-friendliness and communicative relevance of these systems, however far removed they are from established NLP research tracks. Following in the trail of PC-based terminology management and dictionary look-up systems, they form a logical extension to mainstream word-processing technology, essentially the core of the language industries today.

Linguistic Products has recently shipped version 3.1 of its PC-Translator, featuring an extension of its wildcard facility for handling contextual problems of conjugation choice, gender agreement and verb tense for a number of highly-inflected European languages. The system can now handle PageMaker files, permitting document layout integrity during the translation process. A single language pair still costs less than \$1000.

LexiTRANS have recently brought out Version 1.1 of its PC-based *foreign language translation system*, currently offering both-way Russian-English with full Cyrillic script printer support. The company suggests the system is most useful as a means of 'reviewing' foreign language text, perhaps pitching it at information scanners seeking raw versions of technical material in the style of US Airforce use of Systran for technical information scanning.

In a similar register, Barrington Illinois, USA-based Bruce Tolin has brought out version 1 of his eponymous Toltran System. This "revolutionary new patented program," says marketing manager Lori Roberts, can process ASCII files at "up to 1,200 word per minute." This PC-based system sports a "100,000-word database" and comes with 90-days of free technical support via the company's hotline—all for \$2000. Toltran is currently offering an English-Spanish two-way module; alloscript Russian, Chinese and Japanese modules are said to be "under development." No information on non-alphabetic script support is available.

Lacking real user comment, or even reliable test-bed accounts of these systems (although Linguistic Products makes no secret of its user-base), the impression that sales talk is upstaging the growing body of informal industry standards remains strong. For example, MT watchers usually want to know what is meant by the term *word* when judging CAT dictionary size (ie, do we mean headwords, or do we mean conjugated forms?). Figures for raw throughput speed can also be misleading insofar that ancillary activities such as glossary input, post-editing, and document formatting are potentially enormous variables. Language-pair choice can appear to be the result of personal rather than strategic choices. For Europeans confronting real multilingual problems, the systems may reveal more about transatlantic go-gettery than an appreciation of ingrained linguistic habits. There are, after all, no European PC-based translation products on the market. Scepticism perhaps, *eppur si muove*.

Final question: when will practical small-system translationware hit the Macintosh environment?

Linguistic Products, P.O.Box 8263, The Woodlands, TX 77387,
USA. Tel: +1-713 363 9154, Fax: +1-713 298 1911
LexiTRANS, 348 Turnstone Drive, Livermore, CA 94550, USA. Tel: +1-415 447 0596
Toltrans Ltd (International Sales). Tel: +1-617 742 3696, Fax: +1-617 742 3431