A New Era for Systran

By Muriel Vasconcellos and L. Chris Miller

Systran Software's recent release of full,scale machine translation products for the Windows platform marks the beginning of an era. This new product line brings to the desktop the unabridged patented system that has been associated with the Systran name for neatly three decades. It is by far the most massive piece of MT programming ever ported to the PC and offered to the general public. Nine of Systran's twenty-seven language directions are ready to ship: English to Spanish., French, German, Italian, a.nd Portuguese, and Spanish, French, German, and Japanese to English, with Italian and Portuguese in the pipeline. Other language pairs in preparation include Korean, Chinese, and Russian to English.

Systran is known to MT-watchers as the system which the U.S. Air Force has been using since 1969 to monitor technical literature, from Russia during the Cold War, and more recently from France, Germany, Spain, and Japan. The first bidirectional MT software, it won international acclaim in 1975 for its support of the Apollo-Soyuz mission. Systran was also the first MT system to be used by major manufacturers, such as Xerox Corporation, to prepare documentation for products being marketed overseas. Because of computing costs, Systran and its early mainframe siblings were limited to large-scale, high, volume operations. The appearance of a PC-based system of this size and scope brings full,scale MT to a whole new world of users.

System Requirements

A native 32-bit Windows application, Systran Pro runs under Windows 3.1 and is ready for Windows 95 and Windows NT. The stand-alone version requires an Intel 386 processor or higher, a minimum of 16 MB RAM (20 MB recommended), and approximately 15 MB hard disk space per language direction. The price per direction is US$1495. Systran Pro can also be used on a Novell network server in large operations.

Running a Translation

In addition to replicating the strengths of its fore, runner, Systran Pro offers a number of features specially designed for its new platform and its new user-base. These features center around a smooth and snappy interface for handling files, sending and receiving translations, and updating the dictionaries.

Three main windows provide space for the input text, the output text, and optional messages. Input can be in the form of stored files or texts keyed directly into the window. The files may be in WordPerfect, Microsoft Word, RTF, and SGML (Frame and Interleaf optional); Systran Pro will automatically detect the format codes and preserve the page layout, graphics, tables, fonts, and other characteristics of the document. The input stream can be virtually any length, from single words and phrases, to very large docments. Words not found in the dictionary are reproduced in the output text in their original form and at the same time listed in a separate file.

An optional counter in a message window displays how fast the translation is running. We tested several of our own files in six different language combinations (including two versions not yet ready for release). An eight-thousand word English document in WordPerfect with numerous graphics and format codes translated into Spanish in under five minutes on a dedicated 486DX4/100 PC with sixteen MB RAM, resulting in an overall rate of 96,000 words an hour. Four one-page ASCII files, two in Spanish and two in English all translated in twenty seconds or less regardless of the

language direction or whether topical glossaries had been activitated.

Systran Pro can be integrated with the IBM Translation Manager product line and other tools that identify previous translations from a stored database.

Dictionaries

The Systran systems are well-known for their massive, highly refined dictionaries covering a broad range of subject areas. In the Systran Pro product line, each language combination has about half a million unique entries, including idiomatic expressions.

More than twenty topical glossaries come as part of the dictionary and do not need to be installed separately. From an easy-to-use Options menu, the user can select up to four subject areas, either permanently or for a given translation run. The translations in the topical glossaries will override the default translations that would otherwise be generated by the basic dictionary.

Systran Pro enables users to add their own terms and expressions with ease. Also, as a special service, Systran Software will build customer-specific technical glossaries. Because Systran Pro is a true multi-target system, dictionary maintenance is streamlined across languages. In other words, the software lends itself to working from one source language to many target languages simultaneously.

Conclusions

Systran president Denis Gachot takes pride in pointing out that "this entire world-renowned system is now available to almost anyone." The network version of Systran Pro is easily affordable for large-scale users, such as a major in-house translation service or translation bureau. The stand-alone, in turn, will be affordable for smaller translation agencies and possibly for an independent translator with a large volume of business. Systran's entry into the PC market also opens up a host of opportunities besides traditional translation, and it makes full-scale MT available to a new world of users who would otherwise have never been able to afford it.

Systran Software, Inc., 7855 Fay Avenue, Suite 300, La Jolla, CA 92037, USA; Tel + 1 619 459 6700, Fax + 1 619 459 8487, Email info@systranmt.com

Muriel Vasconcellos (71024.123@compuserve.com) is a San Diego,based consultant specializing in translation and machine translation. L. Chris Miller (70303.314@compuserve.com), an MT consultant in Washington, D. C, moonlights as a SYSOP in CompuServe's multilingual WorldForum