

Multilingual word processing - a supplier's view

by Andy Vincent

Vuman Computer Systems have been involved in specialist word processing systems for a number of years, originally with the Vuwriter Scientific package and more recently with the Arts and Language Master versions. In this article we will consider some recent developments in word processing and the particular requirements of multilingual word processing as we see them, and try to predict the developments likely to be of interest to translators and others involved in this area. Inevitably this will be a somewhat personal view, but we hope that it might provide 'food for thought' for people considering buying word processing systems for work with a significant multilingual content.

The extent to which the word processor has become accepted as an everyday tool is probably best illustrated by the success of the Amstrad PCW. That a system produced by a company known almost exclusively for budget price hi-fi equipment should have sold in such numbers to private purchasers and commercial organisations alike now seems unremarkable. However the Amstrad PCW is just a further example of the progress made in the 15 years or so since the first commercial word processors became available. A choice of systems is now usually available for even the most specialised applications. The one exception to this is probably those who require professional multilingual word processing. It is perhaps a symptom of the shortage of good word processors in this area that *Language Monthly* sponsored the Translators' Word Processor project.

The earliest commercial word processors were all dedicated systems - that is computer systems that were only capable of being used as word processors - and in many companies the descendants of these machines still provide the bulk of the general word processing capabilities. However the introduction of the original Wordstar package in the late 1970s provided an alternative solution. This word processor was designed to be just one of many

different programs that could be run on an essentially 'standard' microprocessor-based computer. This approach had, and still has now, the major advantage that it allows users to choose a mix of appropriate software 'packages', often from specialist suppliers, and thus tailor the computer system to their particular requirements. The availability of a wide range of packages for a particular computer enables the hardware manufacturer to expand his market, decrease his costs and ultimately benefits the person buying the computer.

The early 1980s saw a new generation of 'personal computers' using more powerful 16 bit microprocessors and with 16 times the memory capacity of the earlier machines. Greater power allowed the programs to be written in a way that was much more accessible to new or casual users, to become 'user friendly'. Greater memory capacity allowed much more complex programs to be written, and allowed users both more flexibility in the way they used the system, and access to more sophisticated features. Word processor packages for these machines were probably the first to begin to challenge the dominance of increasingly sophisticated dedicated systems.

The greater technical sophistication and inherent cost of 16 bit machines

tended to narrow the variety of machines that became available, consequently software companies were able to begin to more closely tailor their packages to the strengths of particular computers. It was against this background that Vuwriter Scientific was launched at the end of 1982. Vuwriter is analogous to a golfball typewriter equipped with a range of different golfballs. However, unlike the typewriter, with Vuwriter all the 'keyboards' are available simultaneously, and may be instantly interchanged and mixed, and these are combined with all the features of a modern word processor. The other advantage over the typewriter is that the text can be viewed on screen with all the characters correctly displayed, a development known as WYSIWYG (What You See Is What You Get). The Scientific system provides nearly 500 different characters on five keyboards (Normal, Italic, Greek, Scientific, Teletex) chosen to provide a system to meet the requirements of scientists and mathematicians.

The most successful of these newer 16 bit machines has been the IBM PC. The involvement of the world's foremost computer manufacturer signalled that personal computers had reached 'respectability', and has triggered a wave of compatible machines from other manufacturers. These machines tend to offer a

variety of 'improvements' over the IBM original. However, the majority employ one of two types of screen arrangement: Hercules Graphics' for monochrome monitors, and the lower resolution 'IBM Colour Graphics' for colour screens. Recently, these two have been joined by the newer IBM Enhanced Graphics Adaptor which whilst being compatible with the earlier colour adaptor, also supports monochrome monitors at a resolution similar to the Hercules. The higher resolution is generally thought to be less tiring to use and is to be preferred for word processing.

The increasing domination of the IBM PC 'standard' in the mid 1980s gradually forces out all other machines from other than specialist markets. The influx of cheap 'PC clones' from the Far East allowed many users to acquire computers for the first time. The market for word processor packages has grown in step with that for the computers themselves and the packages are becoming increasingly sophisticated and also more specialist. The most successful general word processing packages currently, in no particular order, are Multimate (based on the Wang dedicated word processing system), Word 3 (with multiple font facilities) and updated versions of Wordstar (still the most widely used word processor).

Two further versions of Vuwriter provide an indication of the degree of specialisation now available in packages for the PC. These two versions employ the concepts pioneered in Vuwriter Scientific (now the dominant scientific word processor in the UK and many other European countries) and offer a selection of keyboards targeted at classical scholars (Arts) and linguists (Language Master). The keyboards provided in Vuwriter Arts are Normal, Italic, Classical Greek, the International Phonetic Alphabet and Symbolic. Language Master provides Normal, Italic, Modern Greek, Cyrillic and the characters required in addition to these allow word processing in any other Western European language.

Before drawing some conclusions about the direction of word processing and considering both its future and the criteria for developing a true multilingual word processor, it is useful to consider two parallel developments that are set to converge with word processing. Over the last two years a rapidly growing number of desktop page

printers based on either laser or light emitting diode (LED) technology have become available. The first commercially successful printer of this type was the Hewlett Packard LaserJet, based on Canon photocopier technology; this provided near typeset quality at a price comparable to the cost of a good daisy wheel printer and sheet feed. Since the appearance of this machine a number of manufacturers have introduced their own products, and there are strong signs that this market will develop along the same lines as the PC market, with increasing domination of Far East manufactured printers compatible with the LaserJet, and steadily falling prices. It seems likely that in most applications the laser printer will have replaced the daisy wheel printer within two years, and high quality dot-matrix printers will dominate the lower end of the remaining market.

The second development concerns the increasing number of 'desktop publishing' packages which, when coupled to a suitable laser printer (the Apple Laser Writer was the early leader), allow in-house production of many documents that would normally have required the involvement of graphic artists, typesetters and printers. The advantages claimed for desktop publishing are much greater control over the final appearance, lower cost and greater ease of revision.

The development of the personal computer market has now reached the stage where it is a pre-requisite that a machine is IBM PC compatible, and must be cheaper than those currently available. The latest entry into this market is the AMSTRAD 1512, whose manufacturers expect to sell 100,000 machines a month. The only non IBM PC compatible computer that can be regarded as commercially successful is the Apple Mackintosh and this is sold mainly as part of a desktop publishing package. Consequently, any new word processor package is more likely to be successful if it is targeted at the IBM PC market.

In terms of the development of word processor packages, the situation is not quite as clear cut, but trends are still discernable. The first is that the introduction of low cost entry level versions of the more successful packages has begun to cause dramatic falls in software prices. The second is that the popularity of laser printers has made their support essential, with the inclusion of as many desktop publishing features as

possible. The third trend is that the most successful general purpose package, like the IBM PC itself, are of US origin, and primarily targeted at the US market. Since word processors for this market, unlike those for Europe, do not require significant language capabilities, major packages tend to offer very limited multilingual facilities.

There are two basic criteria which any software package must satisfy if it is to be successful: it must provide the facilities that are required by its intended users and it must be affordable. We will return to the question of cost later, but first we can analyse the facilities required for multilingual word processing in more detail. Talking to any group of word processor users it soon becomes obvious that most people have their own preferred way of working, and will use this as the yardstick by which they judge the suitability of a word processor. It is therefore necessary that the package offers maximum flexibility in the way it is used, and presents its facilities in as simple a manner as possible.

The second major requirement, and the one that is unique to multilingual word processing, is that whilst a particular user may only work in a couple of languages, the total number of languages that users as a whole will require can be quite extensive. Therefore, to maximise the market for a package it must be capable of working with a wide range of languages but not obstruct users who only require a couple. As an extension to this it is also desirable that a file produced in, say German and English, should still be correct when displayed on the machine of a colleague who normally works in Spanish or Italian. In principle this requirement could be satisfied by designing the system to enable the user to configure the system for the particular national keyboard that they require. The complication with this approach is that the preferred method of entering accented characters tends to vary from country to country. In countries with only a few accented characters it is possible to enter an accented character with a single keystroke whilst in others the accent is typed after the characters which it overstrikes (or vice versa). This latter approach is adopted by the IBM PC in European markets but the former is frequently used by existing word processing packages (one notable exception being Vuwriter Language Master). There is an advantage to be gained in storing, but not displaying, the accent

separately in that it offers a significant reduction in the number of characters that must be defined. However, the system should be flexible enough to allow users to operate in the manner appropriate to the languages they are using whilst retaining maximum flexibility.

Further requirements are all refinements of the facilities required in a single language word processor, but they still present some interesting problems. The two most extensively used features of a word processor, after the simple insertion and deletion of text, are block operations and searching/replacing. The basic ability to use block operations to 'cut and paste' text is not affected by multilingual developments. However among other block operations it is also common to be able to force a block of text to become either upper or lower case. In languages where the use of accents is affected by case there are likely to be significant problems. Accents also complicate the normal requirements of search and replace operations, since it is advantageous in languages with floating accents to be able to search for all occurrences of a word without specifying the accents.

One facility increasingly common on word processors is the ability to assign a word or sequence of commands to a single key. This allows complex words or phrases to be entered into the document with a single keystroke. This is likely to be essential on a multilingual word processor and should allow words in any language to be inserted into the document irrespective of the currently selected keyboards.

The final requirement is the availability of an on-line multilingual dictionary. This is likely to be the most difficult to achieve and may remain unattainable for some time to come. The dictionaries on English language word processors are generally a simple list of valid spellings against which the contents of the document are checked. The more sophisticated versions offer an alternative spelling, usually based on phonetic interpretations of any words which they can't locate. However in a multilingual word processor a much more sophisticated tool is necessary so that, in addition to simple checking of spelling, it would allow a word to be entered in one language and offer a choice of translations in a second language. Clearly a wide variety of languages are required. The major problems

with providing a facility of this sort are that the storage capacity that is required is enormous and that it must operate quickly if it is to be a step forward from conventional printed dictionaries. It is likely that this requirement will remain largely unsatisfied until inexpensive CD-ROM systems derived from domestic compact disk players become available.

The second criteria, that a package must be affordable, is likely to be satisfied only if the company producing the multilingual word processor can sell it in sufficiently large numbers, and hence it must be targeted at the only mass market computer, the IBM PC. This does at least have the advantage that the total system cost is also minimised with the easy availability of cheap PC clones.

Finally it is worth considering how close an existing specialist word processor, Vuwriter Language Master, has come to satisfy the requirements of multilingual word processing. This product allows word processing simultaneously in English, French, German, Dutch, Italian, Spanish, Swedish, Finnish, Danish, Norwegian, Icelandic, Polish, Russian, Yugoslavian, Hungarian, Rumanian, Czech and Slovak amongst others. The existing system cannot be configured for particular national keyboards but this is likely to become available in the near future. Satisfying the accent requirements of block case change is a more difficult area and may not be possible until on-line multilingual dictionaries become available, allowing correct spelling to be ascertained.

The latest release of Language Master does, however, provide a powerful and flexible multilingual word processing system including computer and 'letter quality' printer for under £2000. In cases where the output from a low quality printer is acceptable, the system could cost about half this, a price and performance that would have been thought impossible even two years ago. In conclusion whilst multilingual word processing still has some room for improvement, it is possible to buy systems now that offer substantially all of the facilities required by the majority of users.

Andy Vincent, Manchester, October 1986 Dr Vincent has been responsible for the production of the European versions of Vuwriter, and has recently completed work on a Dutch/Russian version for the Dutch Army.