

# Machine translation today

A critical look at current desktop systems by Derek Lewis

The appearance of the first machine translation (MT) system for the PC in the early 1980s marked the beginning of a new phase in MT. From then on, and especially in the following decade, developers began to re-engineer systems designed originally on mainframe-type computers for the potentially much larger desktop market; prices were adjusted accordingly. As a result there are now several systems to choose from, often at prices which compete favourably with standard business applications software. This article discusses some technical advances in MT, especially in relation to the quality of language output, and reviews the types of package currently available. Examples of output from three representative systems are briefly presented in order to give the reader a concrete idea of what to expect from MT. The article concludes with remarks on the difficult task of evaluation.

## Improvements in interfaces

The most obvious advance in MT is the incorporation of windows-based interfaces for the user. These provide rapid and easy access to source and target text files and to system dictionaries. Compared with the facilities of the first MT system for PC, Weidner's DOS-based MicroCat, the latest systems offer clearly laid out, on-screen menus for entering new terms, coupled with extensive on-line help: the ability to update and create dictionaries is particularly useful for users who are willing to customise the system for their own text types and subject domains. In addition, most MT systems are able to maintain original document formatting; some can even be incorporated directly into a wordprocessing package as an on-screen menu option. A far-reaching and very recent innovation is the interfacing of MT with other applications, such as e-mail packages, web browsers, and voice recognition software. In theory, and in the not too distant future, it should be possible to enter a text as direct speech input and listen to the automatically translated results.

## Linguistic capabilities

The main question for professional translators, however, is whether the fundamental linguistic capability of operational systems has improved in line with their interfaces. That this may in fact not be so is suggested by the fact that suppliers on the whole still take care to stress the limitations of raw MT output and are keen not to generate unrealistic expectations on the part of the user. The manual may provide detailed guidelines on preparing text for translation. Examples are: 'use simple sentence structures and punctuation'; 'avoid sentence embedding'; 'rewrite complex sentences into smaller units with a parallel structure'; 'include redundant relative pronouns, prepositions, and other words that clarify sentence structure'; 'avoid strings of prepositional phrases'; 'avoid a nominal style which reduces the role of verbs'; 'introduce relative clauses with a pronoun' (e.g. write: 'the car which was heating the air', not 'the car heating the air ...'). The supplier may even include examples of common error types. This is both to assist in post-editing and to prepare the user for less than perfect quality. Clearly it can be difficult for a supplier to strike a balance between the need to promote a product and the requirement to be honest about its limitations. One supplier illustrates this dilemma when it claims, on the one hand,

that its system possesses a high degree of 'world knowledge' (understood to be extensive grammar rules and dictionary entries) and, on the other hand, includes several pages of advice on how to minimise the complexity of the input text in order to maximise the quality of the output.

It remains the case that MT delivers the best results within a well-defined and carefully managed environment. This generally means using MT for a large volume of texts drawn from limited subject domains in which syntax and lexis are restricted and/or remain relatively stable over long periods. MT is most effectively employed as part of a professionally planned document production and updating process, ideally involving both pre- and post-editing. Indeed some companies already specialise in integrating translation tools within their publishing workflow. This process can involve, for instance, holding multilingual documentation in the form of a layered database in which text units are tagged and structured (e.g. in SGML). Within such a system, individual units are exported for translation and re-imported into the database as required; at the same time standard layouting formats (such as size and length of text, font, etc.) may be applied to particular text units. One of the purposes of managing documents in this way is to preserve consistency of content and presentation across different languages. Properly managed, the approach speeds up the period between the initial writing of a document and its final delivery to the customer (Brougham, 1997).

Until recently it would have been unusual to apply MT to anything but formal, written language. However, the translation of less formal language in the form of on-line texts, notably e-mail messages, is a growth application area that is being specifically targeted by MT suppliers. The potential market for MT here is enormous: the internet generates vast amounts of text across language frontiers, and there is certainly a demand for rapid draft translations where users are more interested in receiving texts for basic information extraction, even at the

Suppliers on the whole still take care to stress the limitations of raw MT output and are keen not to generate unrealistic expectations on the part of the user



ELEPHANT COMPUTER LIBRARY

expense of a loss in stylistic quality. At the same time the linguistic challenges for MT in this context are considerable: like spoken or conversational language, on-line internet 'chat' contains a high percentage of incomplete (ill-formed) sentences, not to mention ellipses, rapid topic shifts, and spelling and grammar errors. Such language is likely to be more difficult to parse than formally correct text. In addition, end users will expect translations to be delivered at little or no cost, although they are likely to be tolerant of output that, while of poor quality by the standards of formal language, provides an adequate basis for mutual communication. CompuServe, who are pioneering MT in this area, have developed a chat translation prototype that translates messages in 5 to 8 seconds; translation of other forms of on-line communication (so-called 'forum messages') takes up to 3 minutes (Fianagan, 1997). CompuServe's aim is to develop an on-line message translation system that is embedded within the e-mail service itself; the MT component, which is based on existing proprietary software, is centrally provided and invisible to the user. Likewise, Altavista have recently made available an on-line MT service for the on-line translation of short texts and web pages; the underlying MT system is Systran and it is available in several language pairs.

Given modern windows environments most MT systems can be interfaced in one way or another with internet applications, a feature which developers and suppliers of established MT systems are now promoting. At its simplest, text can be transferred to and from the MT program via the windows clipboard. MT suppliers who advertise integration with e-mail or the internet include Telegraph and Logos. Transcend's Easy Translator is one system that claims to be designed specifically for translating e-mail and web text.

### Translation Memory systems

Another important advance is the emergence of Translation Memory (TM) systems. The principle behind TM is simple: the computer memorises previous translations and suggests these to the translator as he works on a new source text. The units of memorised translation can be word, phrase, or sentence and can be built up on-line while the user is translating a text or, with the aid of text alignment software, constructed from translations done earlier (so-called 'legacy text'). Typically, a translator works within an already familiar wordprocessing environment, running the TM package alongside it. As the translator moves through the source text, the TM program matches the text with the contents of its database and offers possible translations which can be pasted in directly to the target text; 'fuzzy matches', which

are partial correspondences between source and target text, are also detected and are offered for inclusion or editing, as appropriate.

Although not strictly MT (there is no automatic translation in the sense of analysis and synthesis of source and target language structures) TM has obvious attractions. Firstly, it supports the translator within a familiar work environment. Secondly, it ensures consistency and reliability of terminology, especially for repetitive texts containing small structural units that offer the best chance for successful matching. Thirdly, it allows the translator to draw directly and comprehensively on previous work. Fourthly, unlike MT, TM can be said to 'learn' from texts that have been translated and revised. It can even be integrated with MT programs: in this case the TM system is used to trawl through the source text for matching items before the text is submitted for processing by MT. Finally, as well as exploiting the translator's own previous translations to function as the database memory, it should be possible to incorporate the contents of ready-made, commercially produced dictionaries or term banks: with portability of translation memory databases in mind, there have been recent moves to create standards for the various TM formats (TMX, or translation memory exchange). Suppliers are already making strong claims for TM, reminiscent of early MT. They state that, with TM support, translators can produce up to 12,000 words a day. They stress the time savings achieved and the consistency of translation output. It has even been suggested that users conceal the true benefits of TM in order to enhance their competitiveness (and not to be forced to reduce prices to customers who are unwilling to pay for recycled translations). Suppliers of MT systems include TRADOS, TRANSIT, and T1 (an MT system with a TM module).

### Evaluation

In 1994 an EAGLES subgroup (the EU-supported Expert Advisory Group on Language Engineering Standards) published a draft report on the evaluation of natural language processing systems, ranging from TM and electronic dictionaries to MT proper. Apart from producing a comprehensive

## MT is most effectively employed as part of a professionally planned document production and updating process ideally involving both pre- and post-editing

set of criteria for evaluating translation aids, the report discusses user profiles and summarises developments in the translation industry. According to the report the number of languages in which translations are carried out is increasing. At the same time certain languages are emerging as universal focal languages: this means that a company produces its initial documentation in, say, English, and then translates it directly into other languages. Moreover, while texts for translation may be becoming more standardised and repetitive in nature, translation requires more attention to revision, updating, and layouting; the volume of terminology has also increased. Finally, translations are increasingly contracted to outside organisations; even in-house translation centres are having to operate as independent business units. As for the place of MT in the translation market, the report concludes that MT is still associated with high technology and is to be found in larger, well-resourced organisations with a sophisticated infrastructure capable of providing good support for IT applications.

The question of how to evaluate MT and which system to buy remains a problem, especially for individual translators and small organisations. While laying the foundation of a benchmark-based approach for evaluating translation tools, the EAGLES draft report does not focus on MT; neither does it provide a league table of current systems and their performance. There are in fact a number of methods available for assessing MT and its output, depending on

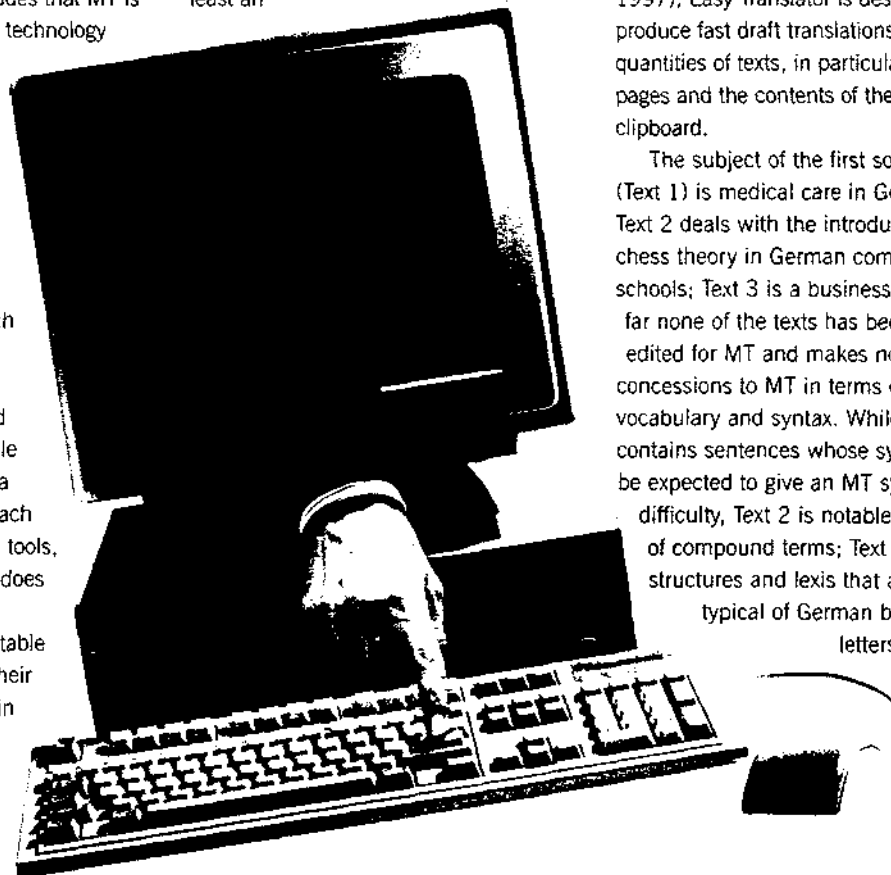
whether you are a systems developer or a user and whether you are applying strictly linguistic or economic criteria. But they must all be applied with care and generally require considerable time and resources to carry through; this usually places them beyond the reach of smaller businesses. For many organisations, the most realistic approach is to sample a system, preferably with the involvement of translators, and, if possible, to consult with existing users. The existence of user groups for specific MT systems would be helpful here, but it is not clear if there are any.

### Examples of output

Below are very small sample outputs from three MT systems. Although it is unwise to draw far-reaching conclusions about a system on the basis of limited input samples, it is possible to gain at least an

initial idea of the basic characteristics of a system; indeed it may be the only feasible approach for an independent translator. In any case a potential user should always subject a system to some kind of test based on the type of text it will be expected to handle. The text samples here are from German to English. The first system is possibly the cheapest currently available and was the first to appear for windows-based PCs: it is the Globalink Power Translator (PT) which translates between English and Spanish, French, German and Russian. The second system is Langenscheidt's T1, a bi-directional English-German package based on METAL which originated at the Linguistics Research Center at the University of Texas in the early 1960s and was acquired by Siemens in 1980. In 1996 T1 for PC was launched as a co-operative venture by Langenscheidt and the Munich-based company GMS (Gesellschaft für Multilinguale Systeme). T1 is a good example of an MT system that has been progressively re-engineered for smaller computers and that has been able to adapt its original model to different language-pairs. The third system is the Easy Translator, marketed by Transcend. A relatively new system (it appeared in 1997), Easy Translator is designed to produce fast draft translations of small quantities of texts, in particular web pages and the contents of the windows clipboard.

The subject of the first source text (Text 1) is medical care in Germany; Text 2 deals with the introduction of chess theory in German comprehensive schools; Text 3 is a business letter. So far none of the texts has been pre-edited for MT and makes no particular concessions to MT in terms of vocabulary and syntax. While Text 1 contains sentences whose syntax might be expected to give an MT system some difficulty, Text 2 is notable for its use of compound terms; Text 3 contains structures and lexis that are fairly typical of German business letters.



<p>German source text 1  <i>Wenn ein durchschnittlicher Patient die bestmögliche Betreuung will, steht er vor einem schwierigen Problem. Offiziell erkundigen kann er sich in Deutschland nicht. Welcher Arzt besonders erfolgreich welche Behandlungen ausführt, ist Standesgeheimnis. Hierzulande sind die Daten von Kliniken und Ärzten für Patienten fast nie zugänglich. Qualitätskontrolle in der Medizin ist ein Tabuthema, obwohl immer mehr Ärzte Transparenz für sich und ihre Patienten wünschen.</i></p>	<p>Output from Power Translator: Text 1  <i>If an average patient wants the optimum care, he/it stands before a difficult problem. Officially he/it inquire can not itself in Germany. Which doctor especially successfully which treatments executes, @@Standesgeheimnis is. Here are the data of clinics and doctors for patients almost never accessible. Quality control in the medicine is a @@Tabuthema, although always more doctor wish transparency for themselves and their patients.</i></p>	<p>Output from T1: Text 1  <i>If an average patient wants the best possible support, he is before a difficult problem. Official can make he itself in Germany not. It is state secret which physician implements which cares especially successfully. The data are from clinics and physicians in these parts for accessible to patient hardly ever. Quality control in the medicine is a taboo topic although more and more physicians request transparency by itself and request their patients.</i></p>	<p>Output from Easy Translator: Text 1  <i>If an average patient wants the bestmögliche care, he stands before a difficult problem. Officially he cannot inquire in Germany. Which physician carries out especially successfully which treatments, is Standesgeheimnis. In this country the data of clinics and physicians are almost never accessible for patients. Quality inspection in the medicine is a Tabuthema although more and more physicians Transparenz desire for itself and its patients.</i></p>
--	---	--	--

<p>German source text 2  <i>Vom kommenden Schuljahr ab können Gesamtschüler Schach als Schulfach in der differenzierten Mittelstufe wählen. Laut Vorschrift soll Schachtheorie, Eröffnungs- und Endspielvarianten zum Unterrichtsstoff gehören. Unklar sei allerdings noch, ob das neue Unterrichtsfach auch in der Oberstufe fortgesetzt werden soll. Auch wer unterrichten soll, steht nach Angaben des Kultusministeriums noch nicht fest.</i></p>	<p>Output from Power Translator: Text 2  <i>From the coming school year from @@Gesamtschüler can choose chess as @@Schulfach in the differentiated junior high school. Loudly instruction should @@Schachtheorie, Inaugural - and @@Endspielvarianten to the subject matter belong. Unclear be indeed still, whether the new subject should be continued also in the higher level. Also who should instruct, not yet is certain after statements of the Ministry of Education and the Arts.</i></p>	<p>Output from T1: Text 2  <i>From the coming school year off, total pupils can select chess as school fold in the differentiated central step. Sound instruction should belong to chess theory, to inauguration and final variants to the instruction material. However, it might still be fuzzy whether the new subject should also be continued in the upper step. Too who should inform is not yet clear according to specifications of the Ministry of Education.</i></p>	<p>Output from Easy Translator: Text 2  <i>Of the coming school year from total students Schach can select as a training specialty in the differentiated means step. According to direction, Schachtheorie, eröffnungs- and Endspielvarianten to the instruction material should belong. Would be unclear to be sure yet whether the new instruction specialty is supposed to be continued also in the waiter step. Also who instruct should, stands according to the Kultusministeriums not yet firmly.</i></p>
---	---	--	--

<p>German source text 3  <i>Wir beziehen uns auf unsere Bestellung Nr. 1 10966, die einer Erfüllungsfrist von 15 Tagen unterlag, die Sie in Ihrem Schreiben vom 28.08. akzeptiert hatten. Bedauerlicherweise müssen wir Ihnen hiermit mitteilen, daß diese Frist Ihrerseits nicht eingehalten worden ist. Die Lieferung war vor einer Woche fällig. Wir sehen uns nun dazu gezwungen, nicht nur den Auftrag völlig zu annullieren, sondern auch von Ihnen Schadenersatz wegen Nichterfüllung des Auftrages zu verlangen. Wir haben unser Rechtsanwaltsbüro entsprechend angewiesen und man wird sich in den nächsten Tagen mit Ihnen in Verbindung setzen. Hochachtungsvoll</i></p>	<p>Output from Power Translator: Text 3  <i>We refer to our order No. 1 10966, which a @@Erfüllungsfrist of 15 days was subject to, which had accepted you in your letter dated 28.08. Regrettably we must inform you herewith, that this time limit has not been observed For their part. The delivery was due before a week. We see ourselves now to it forced, to nullify not only the order fully but also to demand by you damages because of default of the order. We have indicated our law office accordingly and one will sit in the next few days with you in combination. Respectfully</i></p>	<p>Output from T1: Text 3  <i>We refer no to our order. 1 10966, those of a fulfillment period of time of 15 days was subject, those you in your writing of 28.08 accepted had. We must unfortunately hereby inform you that this period of time was not kept for her part. The supply was due a week ago. We now see forced to this ourselves not only to cancel completely the task but also from you compensation because of non-fulfilment of the task demand. We assigned our lawyer's office correspondingly and one will get in touch in the nearest days with you. Hochachtungsvoll</i></p>	<p>Output from Easy Translator: Text 3  <i>We have references to our order no. 1,10966, that was subject to a fulfillment time period of 15 days that you had accepted in your letter of the 28.08., We must announce you with this to Bedauerlicherweise that this time period not on your part was kept. The delivery was due a week ago. We see forced ourselves now in addition of to demand not only the tasks totally of annullieren, but rather also of you damage replacement because of Nichterfüllung of the task. We directed our lawyer office correspondingly and will get in touch one the next days with you in. Respectful</i></p>
---	---	---	--

## An important advance is the emergence of Translation Memory (TM) systems. Suppliers claim that with TM support translators can produce up to 12,000 words a day

### Assessing MT output

What is evident from this output is that MT quality varies quite markedly from system to system: the translations show significant differences in both lexis and syntactic structure. Before considering this issue in more detail, however, there are certain things that we should note about the comparative performance and facilities of our systems, not all of which are evident from just looking at the translations. Speed, for instance, is a case in point. By far the slowest system was T1, which took several seconds longer than the other programs to complete the translations: we would expect the low speed to be offset by markedly superior output quality (the reader may judge for himself whether this is indeed the case). Secondly, T1 allows the user to specify a subject area for an input text: thus, if we tell the system that text 1 is in the domain of medical science, the output changes slightly: the German *Behandlungen* is rendered as 'cares' instead of 'processing'. A final feature of T1 is that it marks words in the target output in various useful ways for possible post-editing. Thus the word 'fuzzy' is highlighted on screen to indicate that there are alternative translations available (viz. 'indistinct', 'unclear', 'vague'); these can be called up on screen and any item from the list pasted in. T1 is also able to break down single-word German noun compounds into their constituents and offer word-for-word translations for them; these are highlighted as potential multi-word terms for post-editing (or entry into the dictionary). In this respect T1 reflects its origins in the METAL system, which was developed specifically for translation between English and German. In our

examples potential terms that have been recognised include: 'state secret' (for *Standesgeheimnis*), 'taboo topic' (*Tabuthema*), 'school fold' (*Schulfach*), and 'central stage' (*Mittelstufe*). Obviously the automatic translation is not always entirely successful, but the advantages over other systems are that a translation is at least attempted (helpful to a post-editor unfamiliar with the source language) and that problems are clearly flagged up for further attention. We should also note that, in addition to the dictionaries accessed by the translation programs, T1 provides on-line bilingual dictionaries that the user can call up in order to review entries whilst post-editing.

Any potential user of MT will be concerned that the output does not fall below a minimum level of comprehensibility. The problem, however, lies in measuring or quantifying comprehensibility. One approach is to assign levels of comprehensibility to each output text, preferably by inviting other people to read the texts and state what they have understood the content to be (e.g. by writing a

summary or answering multiple choice questions). There are at least two drawbacks to this approach. One is that evaluators tend to be subjective in their assessments. Another is that systems often fail to return consistent levels of performance relative to each other. Below are translations of individual sentences (taken from the above texts) ranked in order of comprehensibility (the reader is, of course, at liberty to disagree with the suggested ranking):

German: *Offiziell erkundigen kann er sich in Deutschland nicht.*

ET: 'Officially he cannot inquire in Germany.'

PT: 'Officially he/it inquire can not itself in Germany.'

T1: 'Official can make he itself in Germany not.'

German: *Auch wer unterrichten soll, steht nach Angaben des Kultusministeriums noch nicht fest.*

PT: 'Also who should instruct, not yet is certain after statements of the Ministry of Education and the Arts.'

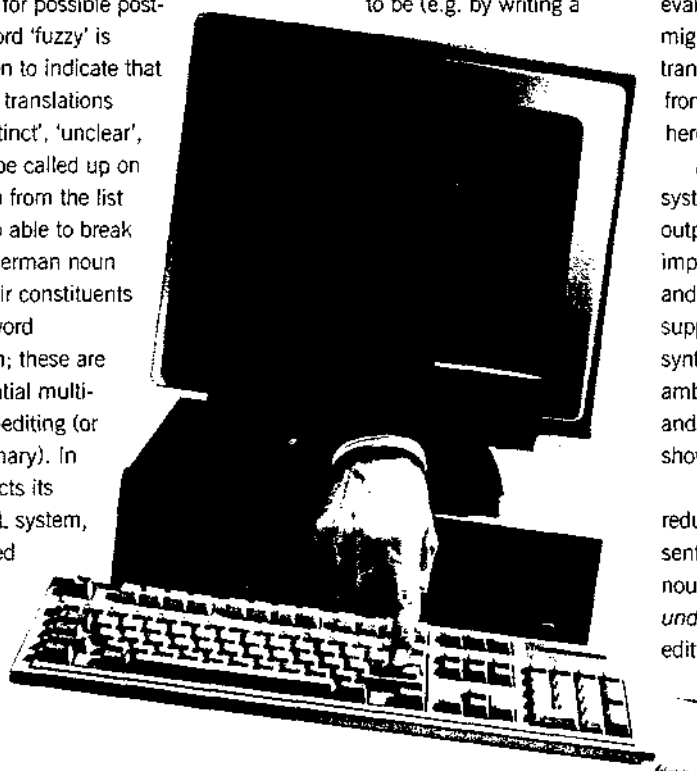
ET: 'Also who instruct should, stands according to the Kultusministeriums not yet firmly.'

T1: 'Too who should inform is not yet clear according to information of the Ministry of Education.'

The point is that the rankings differ for the same systems, although evaluations of large volumes of output might reveal more consistent patterns in translation quality that do not emerge from such a tiny sample as the one here.

A vital factor in assessing any MT system is the degree to which quality of output, however measured, can be improved through dictionary updating and pre-editing of input text. All systems suppliers stress the need to simplify the syntax of input text and avoid ambiguities. Pre-edited versions of Text 1 and Text 2, for instance, might look as shown in the box opposite.

Pre-editing here has been limited to reducing subordination, splitting up sentences and clarifying compounds nouns (e.g. *Eröffnungsspielvarianten und Endspielvarianten*). Effective pre-editing (i.e. of a kind that leads to the best MT output) needs skill and



<p><b>German source text 1: Pre-edited</b> Wenn ein durchschnittlicher Patient die bestmögliche Betreuung will, steht er vor einem schwierigen Problem. Er kann sich in Deutschland nicht offiziell erkundigen. Es ist ein Standesgeheimnis, welcher Arzt besonders erfolgreich welche Behandlungen ausführt. Hierzulande sind die Daten von Kliniken und Ärzten für Patienten fast nie zugänglich. Qualitätskontrolle in der Medizin ist ein Tabuthema. Trotzdem wünschen immer mehr Ärzte Transparenz für sich und ihre Patienten .</p>	<p><b>Power Translator: Text 1: output after pre-editing</b> If an average patient wants the optimum care, he/it stands before a difficult problem. He/it can not inquire in Germany officially. It is a professional secret, which doctor especially successfully which treatments executes. Here are the data of clinics and doctors for patients almost never accessible. Quality control in the medicine is a taboo topic. Still always more doctor wish transparency for themselves and their patients.</p>	<p><b>German source text 2: Pre-edited</b> Vom kommenden Schuljahr ab können Gesamtschüler Schach als Schulfach in der differenzierten Mittelstufe wählen. Laut Vorschrift soll Schachtheorie, Eröffnungsspiel-varianten und End-spielvarianten zum Unterrichtsstoff gehören. Es ist aber noch unklar, ob das neue Unter-richtsfach auch in der Oberstufe fortgesetzt werden soll. Es steht auch noch nicht fest, wer unterrichten soll.</p>	<p><b>Power Translator: Text 2: output after pre-editing</b> From the coming school year from comprehensive school pupils can choose chess as school subject in the differentiated junior high school. According to the regulations chess theory, opening game variants and end game variants should belong to the subject matter. It is still unclear however, whether the new subject should be continued also in the higher level. It not yet is certain also, who should instruct.</p>
---	--	--	--

experience. In Text 2 above, for instance, the editor must decide whether to rewrite the phrase *laut Vorschrift* (e.g. perhaps as *Die Vorschriften verlangen, daß ...* thereby introducing a subordinate clause and causing problems elsewhere in the output), to delete it altogether, or to enter it as a term or phrase in the dictionary. In this case we have decided to store the phrase in the dictionary because it is a standard adverbial phrase and could occur in a variety of text types. Below is the output from the Power Translator after two of the source texts have been pre-edited and the system's dictionary updated.

As a rule pre-editing that concentrates on simplifying sentence structures results in the most consistent improvement in MT output. Customising

a dictionary, for instance, can be a waste of time unless the subject domain is very restricted and the user is certain that alternative translations for certain items will not usually be required. It is also the case that in most MT packages the grammar rules are deeply embedded within the system. As a consequence, the user has little or no access to the syntactic processing rules; individual lexical items are the only things that can be added or modified. As mentioned earlier, many suppliers provide detailed information on the structures to avoid in the input text; such help goes a long way to training users in writing a canonical language that produces the best results for MT. At the same time it is wrong to assume that technical texts operate with a vocabulary and structures that can

entirely eliminate ambiguity and other features of non-technical language.

*Derek Lewis is editor of the Machine Translation Review, published by the Natural Language Translation Specialist Subgroup of the British Computer Society. He is also Director of the Foreign Language Centre in the School of Modern Languages at the University of Exeter.*

**References**

- Brougham, M. (1997) 'Publishing Product Information in a Global Market Place', *Translating and the Computer 19*. Papers from the ASLIB Conference held on 13 and 14 November 1997, London
- Flanagan, M. (1997) 'Online Translation: MT's New Frontier', *Translating and the Computer 19* Papers from the ASLIB Conference held on 13 and 14 November 1997, London