MACHINE TRANSLATION AND COMPUTERIZED TERMINOLOGY SYSTEMS

A TRANSLATOR'S VIEWPOINT.

Peter J. Arthern

Head of English Translation Division,
Council of the European Communities, Brussels.

I have been asked to give a translator's viewpoint on translating and the computer, and I would like to emphasize straightaway that what I am going to say is exactly that - simply a personal impression of the present situation and future developments. While I am fortunate in being able to follow what is going on as a representative of the Council Secretariat on the Commission's "CETIL" Committee (Comité d'experts pour le transfer d'information entre langues européennes) I am not speaking on behalf of the Council Secretariat to-day.

Although I have only a short time available, I want to look at translating and the computer from two points of view. The first is that of a fairly large translating organization which is beginning to use a computerized terminology data base - Eurodicautom - and may become a user of machine translation. The second point of view is that of a translator - and being a staff translator myself I have had to try to put myself into a freelance translator's shoes as well, in order to get a complete picture.

MACHINE TRANSLATION

I am sure the first question which a translator asks about machine translation is "How will it affect my job?" The question was first asked in the 1950s as machine translation projects proliferated in the United States following a demonstration in 1954 by IBM and a research team at Georgetown University. By 1965 American government agencies are estimated to have spent some 20 million dollars in supporting machine translation research at 17 different institutions. And then the Automatic Language Processing Advisory Committee reported in 1966 that machine translation was slower, less accurate and twice as expensive as human translation and that there was no immediate or predictable prospect of useful machine translation.

Whether these criticisms were valid or not, machine translation development in the States was cut back immediately, translators heaved a sigh of relief, and machine translation researchers went underground. As we have already heard this morning however, they are now coming out into the open again and translators are asking the same question once more.

The short answer is that no translator working now is going to lose his or her job in the next five years because of machine translation, and probably never will. Machine translation systems which are now operating are either limited in their scope, such as the Canadian "METEO" system which translates weather forecasts from English into French, or the CULT system which we are to hear about this afternoon, or cannot provide translations of generally acceptable quality without extensive revision, or "post-editing". In ad-, dition, machine translation systems are expensive to develop and can only pay their way by translating large amounts of material. Another bar to using machine translation in small-scale operations is the variety of work, and therefore the variety of terminology involved. If a word is not in the machine's dictionary it just won't be translated, and if a translator has to spend time looking up terms and inserting them in a translation full of gaps, any economic benefit of machine translation will be lost.

Consequently, as things stand at present most freelance translators and staff translators in small firms are unlikely to come into direct contact with machine translation, or to suffer from competition from machine translation.

Competition would only come from the possible use of machine translation by large commercial agencies. It would be felt first either in very general areas, or in very specialized areas, with a clearly delimited vocabulary and standardized phraseology - in both cases, perhaps, in order to have a quick cheap translation to get the gist of a text, or to decide whether to have it translated by a translator.

A final thought in this connection is that both freelances and small firms might conceivably buy raw machine translation from a large agency and post-edit it themselves. This would constitute a particular form of "interactive" machine translation, and would only be worth attempting if the time taken in post-editing to an acceptable standard was less than the time required to translate the text from scratch.

While the size and complexity of machine translation operations mean that freelances and translators in small firms are unlikely to become directly involved with it, some

translators and revisers in the Commission of the European
Communities have already done so.

Some comments on "Systran"

As Professor Sager has told us, the Commission has
bought an American machine-translation system, "Systran",
which it is developing in co-operation with the originator,
Dr. Peter Toma, to translate texts from French into English
and from English into French, and now from English into
Italian. A very small number of translators have been as-
sisting with computer programming in this connection, and
with the preparation of dictionaries. Other linguists have
been revising machine translations into French from English,
in development trials. For them, machine translation has
become a colleague. If it is true that "to understand is to
forgive", this may explain why linguists who have been close-
ly involved in developing the system have been more ready
to revise its output than some others. At any rate, if there
is a future for machine translation in the European Communi-
ty institutions, it is obviously going to involve many
linguists, either as programmers, lexicologists or revisers -
or perhaps "pre-editors".

It is no part of my present responsibility to make an
official assessment of the quality of the Systran trans-
lations already done from English into French at the Com-
mission, but as a matter of interest I asked a number of ex-
perienced colleagues in our own French Division to evaluate
one particular passage from three points of view.

I asked one to read the French translation without
having sight of the English original, and to tell me how
much she understood. This corresponds roughly to the "intel-
ligibility" criterion used in the Commission's own evaluation
of Systran.(1)

I asked a second colleague, the Head of our French Di-
vision, to "mark" the Systran output as if it had been sub-
mitted by a candidate in a competition for recruiting French
translators.

I asked a third colleague - a reviser - to revise the
raw Systran output on the basis of the English original in
exactly the same way as he would revise a translation made
by one of our French translators in the normal course of his
work.

---

(1) Georges Van Slype (1978): "Deuxième évaluation du système
    de traduction automatique SYSTRAN anglais-français de la
    Commission des Communautés européennes". Bureau Marcel
    van Dijk, Brussels.

The last two checks correspond to the "fidelity" criterion also used in the Commission's evaluation of Systran, which most translators will instinctively regard as being the principal criterion applicable in judging machine translation, if not the only one.

The original English text, the raw Systran French translation, and the French text as revised by our reviser are given in Annex I, together with my French colleagues' comments, in the form in which they were kind enough to note them down for me.

Summarizing these comments, the raw translation considered on its own was felt to be quite inadequate for informing a French reader who did not know English about the purposes of the experiment which is described, or the procedure employed. All that he would grasp would be that an experiment with chickens had taken place.

As for the translation considered as an entry in a competition to recruit French translators to the Council's staff, the Head of the French Division wrote that if he had had a paper like this to mark he would probably have stopped halfway down the first page, giving the candidate no marks at all.

The reviser who revised the text on the basis of the English original felt that the machine which had produced the translation has a memory which is far too rudimentary. He considered that evaluation of the system was premature and could not be conclusive because there was no way of assessing the results which might be achieved by a machine equipped with an adequate memory. This was all the more regrettable in that such results provided ammunition for the detractors of machine translation.

Leaving this final conclusion with you, I want now to see how machine translation might affect the operations of the Secretariat of the Council of the European Communities. I must emphasize again that these remarks in no way represent any official position - they simply suggest themselves in the present situation and many if not all will be relevant in any firm or organization which is contemplating the possible use of machine translation.

Possible use of machine-aided translation

Provided that the quality of the final output was acceptable for the purpose in mind, the Council could have three reasons for using a machine translation system. These are, one, that it was cheaper than translation by the traditional translator-reviser system; two, that it was faster, or three, that there were not enough competent linguists

available to produce the necessary translations in any other way.

I do not think there is any possibility of the Council Secretariat using raw machine translation for any of its texts, because the only quality we can accept is 100% fidelity to the meaning of the original, even though the style of our translations - as of many of the original texts - often suffers because of the very short deadlines against which we have to work.

Leaving aside the possible use of machine translation because of a shortage of translators, our principle criterion would be whether we could produce accurate translations faster by using machine translation plus pre-editing and/or post-editing of texts, than we can at present with translators and revisers. In our particular circumstances - where texts have to be sent to the capitals of all the Member States ahead of meetings of the Council, COREPER (Permanent Representatives Committee) and countless working parties - speed is of the essence ana cost is a secondary consideration. Which is not to say that a reduction in costs would not be welcome.

With any given machine translation system, either Systran or the projected European system, we should need to analyse the types of text drafted in the Secretariat solely from the point of view of the total time taken to produce a 100% accurate translation from the original text. Since we could assume that the central processor time in the computer would be identical for every type of text, and could be neglected in comparison with human translation, this would amount to noting the time taken to pre-edit and/or post-edit different types of machine-translated text and comparing this with the total time taken to translate and revise texts of the same type in the normal way. Unless there were clear savings in time, from the moment the original text reached the Translation Department to the moment the completed translation was finally typed, we should not be interested in machine translation.

One would in fact expect that texts which are structured at a superficial level, such as minutes of meetings, would be more amenable to machine translation under our conditions than speeches made by the President of the Council before the European Parliament. Other texts, such as the agendas for meetings, and even the implementing provisions of a Council Regulation or Decision as amended by a working party, can probably be dealt with more efficiently by an extended text-processing system, than by machine translation as such.

In fact, in the Council Secretariat we already employ "translation by photo-copy" to a considerable extent for such things as standard telexes, press releases, appointments to Committees, etc., where it is mainly a question of inserting names, dates, and the titles of documents in a standard format.

Text-processing systems

It is at this point that I go beyond the brief I sketched for myself in the abstract of my paper which you will find in your programme. This is because it has become completely clear to me, since I started preparing for to-day's Seminar, that it is the advent of text-processing systems, not machine translation or even terminology data banks, which is the application of computers which is going to affect professional translators most directly - all of us, freelances and staff translators alike.

Professor Sager and Mr. Tanke have both referred to text-processing systems already, so all I need do is to stress their immense flexibility, in that a single translator working on his own can derive many of the advantages which make a large integrated system so attractive for an organization like the Council or the Commission of the European Communities, with our hundreds of translators.

Some of these are :

- the possibility of amending a text repeatedly on the display screen until it is ready for typing in its final format

- the possibility of changing the layout and presentation of a text, e.g. reducing it from full page width to a single column of half a page width, and typing an amended version alongside

- the possibility of recording standardized chunks of text - paragraphs, whole standard letters, etc. - to be used repeatedly in various combinations

- the possibility of producing texts which look as if they have been set in type, for offset printing.

Extra advantages which the Council, or any large organization with an integrated text-processing system, can expect to derive, lie in the possibility of sending texts from one terminal to another for processing. For example, a text which was being amended during a meeting of the Council could be transmitted page by page to terminals in the various Translation Divisions. The translation of the original text

could be called up on the screen and amended, and the amended translation sent straight back to a terminal with its associated printer in a room next to the Council chamber, so that a complete new text in all required languages could be available by the end of the meeting.

When we were asked at the Council some months ago to co-operate in an enquiry into post-editing systems for machine translation, I commented that our post-editing system consisted of a red ball-point pen in a reviser's hands. This may in fact continue to be true, since if working with a keyboard to revise translated texts at a visual display unit proves to be uncomfortable, it will be a simple matter to have the text printed out and given to a reviser for revision on paper in the traditional way. The corrections will then be made on the text-processing system by a secretary.

This very brief sketch of the possible uses of text-processing systems shows why all translators must consider their use, and also why machine translation systems and computerized terminology data banks must from now on be integrated into text-processing systems.

Further aspects of machine translation

Before I go on to deal with computerized terminology data banks, however, I would like to look at one or two more aspects of machine translation in general.

In a multi-lingual situation such as the one we have in the European Communities, where it is often necessary to produce translations in parallel into several languages from a given original, it will obviously be an attractive proposition - until we get raw machine translation from free-text input which is of almost the same quality as that produced now by our translators - to concentrate on pre-editing texts for machine translation, rather than on post-editing the translations. A good job done on pre-editing a text will save post-editing several translations, and this is a point which those working on the new European machine translation system will presumably have in mind.

If one adopts this approach, however, there will be a tendency to go still further back and to attempt to get the authors of texts to draft them in a standardized form which reduces the need for pre-editing. This is where we come up against resistance - we have already met it in the Council Secretariat when attempts have been made to encourage administrators to use standard formats so as to facilitate translation by photo-copy. And of course, this approach is just not on in the European Parliament or the Economic and

Social Committee, where elected representatives of the people
must obviously be free to express themselves just as they
wish.

If it were possible to dictate to people how they should
write or speak, simply for the sake of making machine trans-
lation cheaper or easier, we could end up by making it more
difficult for them to express themselves in their own
language than it would be for them to learn a second language
and use it.

Finally, there is a real danger that the widespread use
of machine translations which would not be stylistically
acceptable if produced by a translator, even if they convey
the message of the original, would debase and corrupt the
natural languages now in use.

On the other hand, it may be possible in some Community
operations to replace natural language altogether by com-
puterized information. Indeed, only the other day I was
engaged in revising a proposal for a Council Regulation which
contained the following clause: "The documents referred to in
the preceding paragraph or elsewhere in this Regulation may
be replaced by computerized information produced in any form
for the same purpose".

TERMINOLOGY DATA-BANKS

While translators working outside large firms or orga-
nizations are unlikely to come into direct contact with
machine translation, and all translators ought to start look-
ing at the use of text-processing machines or systems imme-
diately, computerized terminology data banks fall between
these two extremes. Their development and use have so far been
restricted to large firms and organizations, but the impend-
ing introduction of publicly accessible data-transmission
networks such as "Euronet" and systems such as "Teletext"
and "Viewdata" which will use the domestic television set as
a visual display unit, may mean that any staff or freelance
translator will be able to dial for information from a term
bank in the not-too-distant future.

Term banks must be user-oriented

We translators may regard machine translation systems as
competitors, and therefore fear them, but we instinctively
feel more at home with something which is obviously not
threatening, since all it can do is to help us in our work.
Term banks must be "user-oriented", as became abundantly
clear at a workshop on "Eurodicautom" which was held in Luxem-
bourg last week, when we found it necessary to spend a con-

siderable time discussing who was intended to use the system, and how, before we could look profitably at its content and structure.

This question of intended use is paramount, since if it is not settled before a system is developed the resulting confusion may be disastrous. In addition to constituting an aid to translation, term banks can also be used for documentary purposes and for standardization - for example, for maintaining single-language normative dictionaries or as mono-lingual or multi-lingual thesauri for information retrieval systems. However, we are only concerned now with bi-lingual or multi-lingual term banks - we can also call them electronic dictionaries - specifically intended to assist translators in the same way as traditional dictionaries.

It will be helpful to extend this comparison, so as to see what a translator can expect from a term bank, and what he cannot. Firstly, he has a right to expect that the information given to him is clearly and logically presented, and can be read easily and quickly. This also applies to normal dictionaries, and is one of the principal criteria normally applied to such dictionaries. Secondly, he has a right to expect that the information given to him is reliable and accurate. However, he must himself decide on the value of this information and make his choice between alternative translations of a given expression, as he does with a normal dictionary.

The basic difference between a printed dictionary and a term bank is that in the term bank all the information is stored electronically and can be added to, updated and amended at will at any time, and that any or all of the information which it contains at a given moment can be made available by a variety of means. It combines the advantages of centralization of information with de-centralization in making it available.

Three ways of presenting information

The information in a term bank can be made available to the translator in three ways; on paper, in the form of a special subject glossary or a text-related glossary; via a television-type screen in a visual display unit used on line; or on micro-fiche used with a micro-fiche reader. The last two ways of looking up information are normally used to answer single queries arising in the course of translating a text, so that the translator will not need to make more than a mental note, or perhaps a hand-written note, of the answer. In both cases, however, it is possible to make a complete record of what appears on the screen, via a printer connected

to the visual display unit or a photo-copier attached to the
micro-fiche reader.

At this point, it will be worth looking at the advant-
ages and disadvantages of all three systems, both for an
individual translator and for an organization using a term
bank.

At the Bundessprachenamt near Cologne, where some 250
linguists are engaged in translating largely technical texts
for the West German Ministry of Defence, a computerized term
bank has been in daily use for the last ten years. The philo-
sophy there has always been to keep the translator away from
the computer and to give him his information on paper, or
on micro-fiche.

The Bundessprachenamt's computer produces two basic
types of glossary. The first is a special-purpose glossary,
printed in a normal type-face, for use by several or many
translators who are all working on a large long-term pro-
ject, perhaps in several places at the same time. The second
is a text-related glossary produced in the form of computer
print-out for a specific text.

In this second case, the translator underlines in his
original text the terms he does not know, or on which he
wants to check, and returns the text to the administrative
office. Here a secretary types these terms into the computer
which prints them out, with their equivalents in the target
language, either in the order in which they appeared in the
original, or in alphabetical order. This list is given to
the translator, who in the meantime has been doing another
job, a few hours later, or the next day.

It is now the translators' responsibility, with the
help of subject codes and other information printed out
alongside the natural language equivalents, to choose
whether the translation offered fits in the context of his
text, and which of a number of equivalents does so, if he
is offered a choice. If the computer offers no translation,
or he is not satisfied with what it provides, the trans-
lator has to find the term he wants by other means open to
all translators, such as looking up normal dictionaries and
reference works, or asking colleagues.

He notes on the computer print-out the new terms which
he finds and uses, and these are then checked by a termino-
logist before being entered in the term bank for further
use, within a fortnight at the latest.

The great advantage of this system for the organization
using it is that it gives constant direct feed-back from the

translators to the system, so that the latest terms are
being recorded all the time and made available to all trans-
lators. In practice, nothing like the same level of feed-
back is produced by the use of visual display units or micro-
fiches.

The advantage of the visual display unit used on line,
both for the translator and for the organization employing
him, is that he can immediately obtain the latest possible
information in reply to a question which crops up while he
is actually doing a translation. This is particularly im-
portant for a Translation Department in an organization like
the Secretariat of the Council of the European Communities,
where many documents have to be translated against very short
deadlines. One can also envisage interpreters consulting such
a visual display unit during a meeting, at least when they are
working in pairs and one interpreter could interrogate the
term bank while his colleague kept talking.

Micro-fiche has the advantage that a very large number
of terms can be stored in a very small space, that it is
cheap to produce, and that it is practicable to distribute the
up-dated contents of a term bank to a large number of users,
both "in-house" and outside the organization, every six months
or so. It would seem at first sight that this might after all
be the cheapest and most practical way of distributing the
contents of tern banks to freelance translators and to staff
translators outside the organizations managing them.

Presentation on visual display units

One important psychological factor in using visual dis-
play units and micro-fiche readers for presenting termino-
logy to translators is that it is not as easy to absorb in-
formation from an illuminated screen as from the printed
page. If a term bank is designed for use by either of these
methods, it is vital that the information which the trans-
lator wants should be presented to him clearly in a minimum
of words, and without any unnecessary visual clutter.

This point in fact is so important that it really means
that the presentation of information in a term bank which is
going to be used on line at all must be designed for this
purpose. If the presentation is acceptable on the screen, it
should be completely acceptable on paper, but the reverse
is not true.

In order to give some idea of the practical considerat-
ions involved in consulting a term bank on line from a vi-
sual display unit I should like to describe my experience
operating a terminal installed in the Council Secretariat
in Brussels, and connected via a dedicated telephone line

to "Eurodicautom", the Commission's terminology data base
at the Computer Centre in Luxembourg.

First of all, it is obvious that the technology at
present being used for long-distance connections is not yet
satisfactory, as there are fairly frequent disturbances and
interruptions to the service for technical reasons. For
example, during a recent two-hour session at the terminal,
it was only possible to interrogate the term bank for about
two thirds of the time during which the terminal was con-
nected.

The actual operation of the terminal is very simple and
it only requires half an hour or so to grasp the mechanical
operations involved, many of which are simplified by the pro-
vision of special keys for commanding various functions,
such as asking a new question, or a decision to operate the
truncation of the words requested - of which more later -
or to have the associated printer print out the text ap-
pearing on the screen.

What does require a little practice and - until an
operating handbook is available - experimentation, is to
discover the optimum way of putting questions in order to
get the most helpful answer as quickly as possible. This is
because the system is designed to give partial information
in reply to a question when it does not contain an equiva-
lent for the whole expression which has been requested, and
the user can get bogged down in a mass of irrelevant answers.

A question is put by typing on the keyboard the term or
expression for which the correct equivalent in the target
language is wanted. As the words are typed, they appear on
the screen. When the operator has checked that the expression
appearing on the screen is correct, he presses a special
"enter" key to the right of the space bar and waits for the
answer to come up on the screen. If the first answer is not
completely satisfactory, further answers, each reproducing
the content of a distinct entry in the "dictionary", can be
called up by pressing the "entry" key again after each suc-
cessive answer. When there are no more answers relating in
any way to the question which has been put, a message to
this effect appears on the screen.

Articles or prepositions which appear in the "question"
should not be typed, since the system neglects them unless,
as is the case with the French preposition "de", confusion
is possible (accents not being taken into account) with
nouns. In such a case, typing a preposition can call up
false answers, and so slow down the operation.

On the principle of the longest match, the system will

normally give the correct answer to an expression containing three or four significant words as the first answer, if it contains the expression as such at all. If it does not, one should press the "truncation" key at once, because this will produce the answer if any word or words in the question were in the singular while they are in the plural in the expression recorded in the term bank or vice versa. Even with an expression containing only two significant words, dual or multiple meanings are rare, so that if the term bank contains the answer one is looking for, it will usually come up as the first one.

The difficulty starts when one has entered an expression containing more than one significant word, for which the system has no exact match. In this case, in an effort to be helpful, it looks through its memory for any occurrence of any of the single words in the expression, and at present brings them out in an apparently random order, depending on the chronological order of their entry into the term bank.

The same random plethora of information is liable to appear when one enters a question consisting of a single word, particularly if it is a common one. But perhaps one should not be asking Eurodicautom simple words?

Be that as it may, I have found in practice that if the answer one wants does not appear as the first or second (after truncation) answer to the question, it is rarely worth continuing to press the "enter" key to obtain more than five answers. For this reason, and because it takes the printer one minute and five seconds to print a screen full of information, and it cannot be stopped at the end of the actual text on the screen, so that it may be "printing" empty space for half its time, I have designed a reply form which I use to note relevant information long-hand. This form is shown in Annex II.

If one's answer comes up first time, I have found that one obtains it in between 15 and 45 seconds after starting to type the question. As this time includes typing, it obviously depends on the length of the question, and I am only a two-finger typist, so experienced operators will obviously be able to do better. To write out the relevant parts of five answers long-hand in completing one of the special reply forms takes an average of three minutes.

"Feeding" a term bank

Having spent some time in looking at how information can be obtained from a term bank - as this obviously affects translators who are using it - it will be as well to examine how information should be put into it, and by whom.

It would be technically possible to allow any user who had access to a visual display unit with keyboard to add new material, or to amend what was already recorded. This is obviously undesirable, but it is equally undesirable to exclude users from contributing to the term bank at all, since the most fruitful way of running a term bank is to have a constant symbiosis or "osmosis" between users and the terminologists who are responsible for what goes in.

The principle here must be that users are positively encouraged to submit proposals at all times, either for the translation of expressions which they have not found in the system, or because their experience tells them that their suggestions may be useful. Of course, these proposals must be vetted by the terminologists before they are entered, but this should be done within a fortnight of the proposal being submitted, as experience in systems operating in this way shows that translators want to be able to check that their proposals are in the system within this time, otherwise they become discouraged.

This collecting of terms at the "front line" of translation can of course be backed up by systematic research by professional terminologists in areas which it is felt the term bank should cover, using all the traditional tools and methods, such as reading original specialized texts in all the languages in which one is interested. It is also possible in a large organization which is running a term bank to set up ad-hoc mixed teams of terminologists and translators or revisers to collect expressions in all the relevant languages in a particular field in which they are working. We have done this at the Council Secretariat, for example, to produce a six-language glossary of all the working parties and other bodies operating under the auspices of the Council.

Whichever method is used, speed in getting the results into the term bank is of the essence, particularly where one has a large number of translators working on important texts against urgent deadlines. The only acceptable method is now the use of keyboards keying directly into the memory, as in the commercially available text-processing systems. And if it is true, as I saw yesterday in someone else's newspaper, that it is now possible, in principle, to store half a million pages of text on a single memory disc, all of it immediately accessible, we shall have a very simple method of instantly amending and updating very large term banks.

Organizations which have already set up term banks, or which are contemplating doing so, will have made their decisions for a variety of reasons, not all of which will be relevant to a freelance translator or a staff translator in a small firm. However, the advent of increasingly flexible

text-processing systems will mean that many small firms may find it worth using their typing equipment in order to set up a private tern bank on the side.

Is there a market for term banks?

What, though, is the market going to be for selling terminology from a term bank to independent "outside" trans- lators, either freelance or staff? If anyone is contemplating doing this, he should do some hard market research first, because people are not going to keep on paying in order to find out, after dialling a term bank, that it doesn't con- tain the answer they want.

I have emphasized dialling for information, i.e. inter- rogating a term bank on line via "Euronet" or "Viewdata" etc., because this is the only really new development in making information available, with the one prime advantage over the printed word that the information can be constantly up-dated without it being necessary to send subscribers loose- leaf addenda or printed supplements to the main body of a glossary. Translators who buy the output from a term bank in the form of printed glossaries or micro-fiches will obvious- ly judge it as they judge a dictionary. They will have paid for their information in advance, probably on the recommend- ation of colleagues or of professional publications. Their decision as to whether they have got their money's worth cannot cancel their original purchase; at best (or worst) it can only determine whether they place a repeat order or con- tinue their subscription.

I imagine that an outside subscriber dialling for instant information from a term bank would be charged for every call he put through, whether or not he found the answer to his question. And even if the service was free, he would not con- tinue dialling if he did not obtain a high proportion of satisfactory answers.

In addition to clear presentation of the information they contain, the second essential requirement for term banks designed to be used on line by translators is therefore that they give their users a sufficiently high ratio of satisfac- tory answers. This criterion applies both to in-house staff in a large organization and to outside subscribers. Possibly one group would accept a lower ratio of satisfactory answers than the other.

Co-operation between term banks

This need to provide a high ratio of answers has led the managers of existing term banks to look at ways of exchanging

information between term banks. "Eurodicautom" has been ac-
tive in this area, and an ISO working party has been study-
ing possible standards for the exchange of data on magnetic
tape. Experience *so* far seems to indicate that the difficul-
ties in the way of exchanging information are in the main
not technical (incompatibility between computer programs and
equipment), as was at first thought, but managerial, in the
sense that differing term banks have different philosophies
and different ways of presenting information, so that in-
formation from outside has first to be checked against what
is already in the system, in order to prevent duplication,
and then tailored to fit.

There is a second drawback to the simple exchange of in-
formation between term banks in that it will, if carried to
its logical conclusion, lead to the existence of several
identical term banks all containing the same information.
This would at least make it easier for the independent trans-
lator - he would simply dial his local term bank, instead of
having to find out by trial and error which one gave him the
best service.

The logical solution is surely that term banks should
continue to be set up wherever they meet a particular local
need, and that all of them should pass on the terminology
which they record to a central term bank for a particular
geographical and/or linguistic area. These central term banks
would themselves be linked to a single world term bank, pre-
sumably under United Nations management. Unfortunately, per-
haps because of financial limitations, the existing UN ter-
minology body, "Infoterm" in Vienna, is not pursuing this
line of thought.

How, then, should bi-lingual or multi-lingual term banks
develop in future so that all linguists can make optimum
use of them? As I have already hinted, questions of intended
use and presentation of information are much more fundamental
than the data-processing techniques employed, although tech-
nical incompatibilities should certainly also be reduced to a
minimum.

A point which needs to be re-emphasized here is that
there are in fact distinct classes of potential users of such
term banks, i.e. translators (and revisers), terminologists
and documentalists. Translators usually simply want to have
the correct equivalent for a term or expression which they do
not know, or on which they want to check, accompanied by a
note in plain words indicating context, usage etc., if the
term is not self-explanatory. Terminologists and documenta-
lists however need more information, and experience is show-
ing (currently in Denmark) that it is impossible to present
all the information which they want on a visual display unit

without cluttering up the screen unnecessarily for transla-
tors.

Two-stage presentation

I therefore propose that a standard two-stage format for
the presentation of terminology on the screens of visual dis-
play units should be agreed internationally as soon as pos-
sible. When a user first keyed in a question, he would receive
only the "translator package" of information. Terminologists,
documentalists and even curious translators could then re-
ceive the supplementary information, such as source, defini-
tion, illustrative context, subject codes, etc., presented
below the first basic package on the screen, by pressing a
second key on the terminal keyboard.

Consideration should also be given to presenting a series
of "translators' packages" on the screen simultaneously, one
below the other, so that the screen would read like a page in
a well-designed glossary. Since experienced translators can
very quickly scan a whole page of a glossary or word list,
this form of presentation, avoiding the need to key in for
successive entries which appear on the screen one at a time,
would speed up the process of interrogation very considerably.

If everyone operating a term bank, however small, were
to use this standard format for presenting their information,
allied with strict respect for technical standards for trans-
ferring information between term banks on magnetic tape,
floppy discs, or other forms of memory yet to be developed,
this would be a giant step towards the centralizing of
terminology records for which I have already put in a plea.
It would also mean that everyone would quickly learn to use
information from any term bank, since the technique of inter-
rogation would be the same for all of them.

In this crystal-gazing exercise, I have concentrated on
access via visual display units, but it seems to me that
standardization of presentation would also have advantages
for micro-fiches and printed glossaries. The layout of the
latter could in any case be varied at will to meet particu-
lar requirements by the use of standard text-processing
techniques as now applied to typed and printed documents.

TRANSLATION BY TEXT-RETRIEVAL

Having looked at machine translation and terminology
data banks separately, with brief references to text-proces-
sing systems, I now want to sketch further possible develop-
ments based on such systems.

In the first place, it has become evident during the

Systran trials already carried out by the Commission of the
European Communities that machine translation makes no sense
unless it can be fitted into the normal production line for
translations. As the obvious way of entering, pre-editing
and post-editing machine translation texts is now to use a
text-processing system, this has led to the realization that
the whole production process for translations in the European
Community institutions should be re-designed so as to make
the maximum use of all the potentialities of large text-pro-
cessing systems, whether or not machine translation as such
is ever used on a routine basis or not.

"Controlled" situations

From this realization it is a short step to the pro-
posal which I now put forward for a new form of machine-aided
translation which could give immense benefits in a large
"controlled-translation" situation such as that existing in
the European Community institutions. In the Community insti-
tutions a large number of linguists are employed to translate
enormous amounts of written text, in a variety of original
languages, into several languages simultaneously. In addition,
and this is equally important, all these texts refer to a
"controlled" situation, in that the field to which they relate,
although very wide, is not infinite, and could in theory be
precisely defined at any given moment. Finally, many of the
texts involved are highly repetitive, frequently quoting
whole passages from existing Community documents.

If, as frequently happens, authors do not indicate the
source for their quotations, it is easy to imagine how much
time is quite unnecessarily wasted by translators in search-
ing for references, or in re-translating texts which have
already been translated.

Many of these characteristics, if not all, will also
be present in other international bodies, government depart-
ments and industrial and commercial undertakings. If such
bodies are looking at the use of text-processing systems
for handling their normal documentation and correspondence,
they might also consider their potentialities for dealing,
as follows, with their translation problems.

All texts stored in a single memory

The pre-requisite for implementing my proposal is that
the text-processing system should have a large enough central
memory store. If this is available, the proposal is simply
that the organization in question should store all the texts
it produces in the system's memory, together with their
translations into however many languages are required.

This information would have to be stored in such a way that any given portion of text in any of the languages involved can be located immediately, simply from the configuration of the words, without any intermediate coding, together with its translation into any or all of the other languages which the organization employs.

This would mean that, simply by entering the final version of a text for printing, as prepared on the screen at the keyboard terminal, and indicating in which languages translations were required, the system would be instructed to compare the new text, probably sentence by sentence, with all the previously recorded texts prepared in the organization in that language, and to print out the nearest available equivalent for each sentence in all the target languages at the same time, on different printers.

Grammatically correct partial translations

The result would be a complete text in the original language, plus at least partial translations in as many languages as were required, all grammatically correct as far as they went and all available simultaneously. Depending on how much of the new original was already in store, the subsequent work on the target language texts would range from the insertion of names and dates in standard letters, through light welding at the seams between discrete passages, to the translation of large passages of new text with the aid of a term bank based on the organization's past usage.

When the completed translations were typed in the processing system, they would at the same time be entered in the text memory in association with the original, so that the store of translated texts would be automatically updated.

Further considerations

The texts stored in this way could also be used as a source of "raw" terminology by calling up individual words or expressions on the screen, with their equivalents in other languages. Terminologists would check and process this information in order to enter it in a separate term bank memory in the internationally agreed format, but if a translator wanted a particular term before it was in the term bank, he could look it up in the text store.

Since this form of machine-assisted translation would operate in the context of a complete text-processing system, it could very conveniently be supplemented by "genuine" machine translation, perhaps to translate the missing areas in texts retrieved from the text memory. Whether these mis-

sing areas were translated by translators, or by a machine,
the terminology used would have to be identical, and must be
consistent with the normal terminology employed by the orga-
nization.

This latter aspect of machine-aided translation has al-
ready cropped up in the European Communities, where I and
others have been urging for some years now that the machine
dictionaries used for the Systran trials should be consistent
with the information contained in "Eurodicautom". Those
working on these two projects in the Commission are well
aware of this requirement, but the same type of considerations
apply here as in the exchange of terms between term banks,
with the added complication that a machine translation dict-
ionary has to contain vastly more coded information than a
term bank for translators or terminologists.

THE TRANSLATION BUREAU OF TOMORROW

Pulling all the scattered aspects of my paper together,
what will it be like to work as a translator/reviser/post-
editor in the computerized translation bureau or department
of tomorrow? Do not forget either that, given reliable tele-
communications, a freelance translator will be able to have
all the facilities at home which his staff colleague will
have at the office.

My hunch is that our translator - in many cases, we
ourselves - will continue to work at the same type of desk
in the same type of office which he (or she) has to-day,
with his standard dictionaries and reference works around
him. Instead of a traditional type-writer, however, he will
have a text-processing terminal with keyboard and screen so
that he, or a secretary to whom he dictates, types his
translations into the system memory so that they can be
corrected on the screen before final "typing" on a separate
printer which he will share with a number of colleagues,
unless he is working as a lone freelance.

If he has access to a local term bank, he will be able
to interrogate it simply by typing his question on the key-
board of his text-processing terminal, when the answer will
appear on the screen and can also be printed out by the
printer. It will also be possible for him or his secretary
to get a text-related glossary from the term bank, via the
printer, by using the terminal to type questions into a
buffer memory for batch processing.

In a large organization using my proposed new system
of machine-aided "translation by text-retrieval" (let's call
it "TERRIER" - an appropriate name, since the Shorter Oxford
Dictionary defines this word as "an inventory of property

or goods" as well as "a small, active, intelligent variety of dog which pursues its quarry into its burrow or earth") our translator will be given, when he reports for duty, not only the original of the text he is required to "process", but TERRIER'S version of it in the target language, which we hope will be his mother tongue, both presented on paper in normal type-script.

Secure in the knowledge that he does not have to do any research for possible hidden references, since TERRIER has done this for him, except for references to documents not already in the system, he will complete the target-language version of the text on paper, using his text-processing terminal to type any completely new passages. He will also use his terminal to get terminological information from the organization's term bank if necessary, either on line or in the form of a text-related glossary if he has enough time.

He will then check the complete translation and pass it on, either for revision, if a separate revision stage is re-quired, or straight for typing by a secretary into the text-processing system for storage in the text-memory and printing out in whatever form is required.

It would of course be technically possible to do all translating, editing and revision operations on the screen at the terminal, without printing the texts on paper at all, but I rather suspect that, except for extremely urgent or fairly simple texts, people will prefer to continue getting at least the final versions of their work onto paper so that they can carry out a final check, or so that a reviser can revise the text, with the good old-fashioned pen, pencil or ball-point, unharassed by modern technology.

After all, translation is in the end a creative activity, not a mechanical chore.

<u>ANNEX I</u>

**COMMENTS ON A TEXT TRANSLATED
FROM ENGLISH INTO FRENCH BY THE
COMMISSION'S "SYSTRAN" MACHINE
TRANSLATION SYSTEM DURING THE
SECOND EVALUATION (1978)**

<u>THE ENGLISH ORIGINAL</u>

<u>Introduction</u>

Before final implementation of Regulation (EEC) 2967/76 of
23rd November, 1976 on the water content of frozen and deep-
frozen chickens, hens and cocks the question of the corre-
lation between the two analytical methods for determining
extraneous water in that regulation has been raised.

In order to compare the two analytical methods (annex III,
the protein method, and annex IV the fat-free dry matter-
method) the services of the EEC Commission asked two labo-
ratories - Bundesanstalt für Fleischforschung, Kulmbach,
Germany and the Danish Research Institute for Poultry Pro-
cessing, Hillerød, Denmark - to carry out comparative ana-
lyses on deep-frozen chicken and hen carcasses with the two
methods.

Assisting in planning and carrying out the study were also
Dr. P. Stevens, Station de Recherches Avicoles - INRA,
Nouzilly, France, and Dr. J. van Hoof, Laboratorium voor
Hygiene en Technologie van Eetwaren van dierlijke Oorsprong,
Gent, Belgium.

The study was numbered P.200 by the EEC Commission services,
and the actual work of the study was divided by the labora-
tories in the following way: All hens were slaughtered by
the Gent Laboratory, and the hens were divided between
Kulmbach and Hillerød. All chickens were slaughtered by the
Hillerød institute and they were divided between Kulmbach
and Hillerød.

This report describes the Danish part of the study for the
chickens analyzed in Denmark. A report on the Danish analyses
of the hens will form part II of this report.

Live chickens. The chickens for the study were taken from
two small flocks of the same genetic origin reared in the
Research Station for Poultry Breeding at Hillerød. They were

at the time of slaughter (17th and 18th October, 1977) 39-40
days and 55-56 days respectively. They were White Plymouth
Rocks. They have been given a feed with the following com-
position:

40% maize
20% barley
24% toasted soy bean cakes
 3% meat and bone-meal
 4% fish meal
 3% oats
 3% animal fat
 3% minerals and vitamins

All had been starved for at least 12 hours before slaughter.
On each of the two days 3 times 25 chickens (for three
weight groups) were caught and transported to the near-by
slaughterhouse (distance app. 500 meter). Each weightgroup
should contain 24 chickens - the 25th being an extra, that
was slaughtered but not used in the study.

Slaughtered chickens. All chickens were electrically stunned
in a water stunner, killed and bleeded, but removed from the
slaughter-line before the scalding tank. They were then hand-
plucked and eviscerated. All organs and the neck were re-
moved. They were dressed according to the definition:
"plucked and drawn, without heads and feet, and without
hearts, livers and gizzards, called 65% chickens", and with
the removal of the neck.

The carcasses were then divided in two groups (12 from each
weight class) and numbered by application of rubber rings
around the thighs, the same number being applied to the two
thighs of the same chickens. Those destined for immediate
deep-freezing were given a number preceded by a Roman I,
those destined for wet chilling with a number preceded by
a Roman II.

THE RAW SYSTRAN OUTPUT

Introduction

Avant l'exécution finale du règlement (la CEE) 2967/76 de
23rd novembre 1976 sur la teneur en eau des poulets, ces
poules et des coqs gelés et surgelés la question de la cor-
rélation entre les deux méthodes analytiques pour détermi-
ner de l'eau étrangère à ce règlement a été soulevée.
Pour comparer les deux méthodes analytiques (annexe III, la
méthode de protéines, et l'annexe IV la import-méthode sèche
non grasse) les services de la Commission de la CEE ont de-
mandé à deux laboratoires - la fourrure Fleischforschung
De Bundesanstalt, Kulmbach, l'Allemagne et l'institut de

recherche danois pour la volaille traitant, Hillerod, Dane-
mark - d'effectuer des analyses comparatives sur profond -
des carcasses gelées de poule et de poulet avec les deux
méthodes.

Aider à la planification et en effectuant l'étude étaient
également dr. P. Stevens, poste De Recherches Avicoles -
INRA, Nouzilly, France, et le sabot de dr. J. van, le camion
anglais dierlijke Oorsprong, Gent, Belgique De Eetwaren de
camion De Technologie de l'hygiène de Laboratorium voor.

L'étude a été numérotée P 200 par les services de la CEE de
la Commission, et le travail réel ce l'étude a été divisé
par les laboratoires de la manière suivante : toutes les
poules ont été abattues par le laboratoire De Gent, et les
poules ont été divisées entre Kulmbach et Hillerod. Tous les
poulets ont été abattus par l'institut De Hillerod et ils ont
été divisés entre Kulmbach et Hillerod.

Ce rapport décrit la partie danoise de l'étude pour les pou-
lets analysés au Danemark. Un rapport sur les analyses dan-
oises des poules fera partie II de ce rapport.

Poulets vifs. Les poulets pour l'étude ont été pris de deux
petits troupeaux de même origine génétique élevés dans le
centre de recherches pour volaille élevant au Hillerod. Ils
étaient lors octobre d'abattage (de dix-septième et dix-
huitième, de 1977) 39-40 jours et de 55-56 jours respective-
ment. Ils étaient les rochers blancs Plymouth. Ils ont été
donnés un fourrage avec la composition suivante :

maïs de 40 %
orge de 20 %
24 % ont grillé des gâteaux de haricot de soja
3 % viande et engrais d'os
farine de poissons de 4 %
avoine de 3 %
matière grasse animale de 3 %
minéraux et vitamines de 3 %

Tout avait été affamé pendant au moins 12 heures avant
l'abattage. Sur chacun ces deux jours 3 fois 25 poulets
(pour trois groupes de poids). Être recueilli et transporté
à l'abattoir voisin (distance app. 500 compteurs). Chaque
poids-groupe doit contenir 24 poulets - 25th être un figu-
rant, qui a été abattu mais pas n'a pas été employé dans
l'étude.

Poulets abattus. Tous les poulets ont été électriquement
assomés dans une eau stunner, tués et bleeded, mais enlevés
de la abattage ligne avant le réservoir d'échaudage. Ils
hand-plucked et ont été éviscérés alors. Tous les organes

et le col ont été enlevés. Ils ont été habillés selon la définition "plume et tiré, sans ces têtes et des pieds, et sans des coeurs, foies et gésiers, appelés poulets de 65%, et avec l'élimination du col.

Les carcasses ont été alors divisées dans deux groupes (12 de chaque classe de poids) et numérotées par l'application des anneaux en caoutchouc autour des cuisses, le même nombre étant appliquée aux deux cuisses des mêmes poulets. On a donné ces destiné à la surgélation immédiate un nombre précédé par un Roman II.

THE SYSTRAN OUTPUT REVISED

Introduction

Avant la mise en oeuvre finale du règlement (CEE) n° 2967/76 du 23 novembre 1976 sur la teneur en eau des coqs, poules et poulets congelés ou surgelés, la question de la corrélation entre les deux méthodes d'analyse proposées dans ce règlement pour déterminer la teneur en eau étrangère a été soulevée .

Pour comparer les deux méthodes d'analyse (annexe III, méthode des protéines, et annexe IV, méthode des matières sèches non grasses), les services de la Commission de la CEE ont demandé à deux laboratoires - la Bundesanstalt für Fleischforschung de Kulmbach, Allemagne, et le Danish Research Institute for Poultry Processing de Hillerød, Danemark - d'effectuer des analyses comparatives sur des carcasses surgelées de poulets et de poules suivant les deux méthodes.

Ont également contribué à programmer et à réaliser cette étude le dr. P. Stevens, Station de Recherches Avicoles - INRA, Nouzilly, France et le dr. J. van Hoof, Laboratorium voor Hygiene en Technologie van Eetwaren van dierlijke Oorsprong, Gand, Belgique.

Les services de la Commission de la CEE ont donné à cette étude le numéro P 200 et les laboratoires ont réparti de la manière suivante les travaux à effectuer pour réaliser celle-ci: toutes les poules ont été abattues au laboratoire de Gand et elles ont été réparties entre Kulmbach et Hillerød. Tous les poulets ont été abattus à l'institut de Hillerød et ils ont été répartis entre Kulmbach et Hillerød.

Le présent rapport décrit la partie danoise de l'étude effectuée sur les poulets analysés au Danemark. Un rapport sur les analyses réalisées au Danemark sur les poules fera l'objet de la partie II de ce rapport.

Poulets vivants. Les poulets qui ont fait l'objet de cette étude ont été choisis dans deux petites basses-cours de même origine génétique et ils avaient été élevés au centre de recherches pour l'élevage de la volaille de Hillerød. Lorsqu' ils ont été abattus (les 17 et 18 octobre 1977) ils avaient respectivement 39-40 jours et 55-56 jours. Il s'agissait de White Plymouth Rocks. On leur avait donné un fourrage dont la composition était la suivante :

maïs: 40%
orge : 20%
tourteaux de soja grillé : 24%
viande et engrais d'os: 3%
farine de poisson: 4%
avoine: 3%
graisses animales: 3%
minéraux et vitamines: 3%

Ils avaient tous été privés de fourrage pendant 12 heures au moins avant d'être abattus. Pendant chacun de ces deux jours, 3 lots de 25 poulets (correspondant à trois catégories de au poids) ont été constitués et transportés à l'abattoir qui se trouvait à proximité 500 mètres environ. Chaque lot devait comporter 24 poulets; le 25ème étant en réserve, a été abattu, mais il n'en a pas été tenu compte dans l'étude.

Poulets abattus. Tous les poulets ont été assommés électriquement dans un réservoir d'eau, tués et saignés, mais retirés de la ligne d'abattage avant le bac d'échaudage. Ils ont été plumés à la main, puis éviscérés. Tous les organes et le cou ont été enlevés. Ils ont été préparés selon la formule "plumés et parés, sans les abats (tête, pattes, coeur, foie et gésier), appelés poulets 65%", et sans cou.

Les carcasses ont alors été réparties en deux groupes (12 de chaque catégorie de poids); par ailleurs, on les a numérotées au moyen d'anneaux en caoutchouc qui leur ont été passés autour des cuisses, le même nombre figurant sur les deux cuisses d'un même poulet. Pour celles destinées à la surgélation immédiate, ce nombre était précédé d'un I romain; pour celles destinées à la réfrigération humide, celui-ci était précédé d'un II romain.

<u>COMMENTS</u>

1. <u>The French text considered on its own</u>.

(1)   Pour un lecteur francophone non prévenu qui ne connaîtrait aucune langue étrangère, l'ensemble du texte apparaît presque incompréhensible en raison des nombreux non-sens, des erreurs de construction, des mots non traduits, etc. Il est probable qu'après avoir achevé sa lec-

ture, ce profane ne pourrait rien dire ni de la finalité, ni des modalités de l'expérience. Tout ce qu'il saurait, c'est qu'une expérience a eu lieu. Autrement dit, l'information retirée serait pratiquement nulle.

(2)   Il est vraisemblable qu'un lecteur déjà au fait de l'expérience, surtout s'il connaissait déjà l'anglais (et le néerlandais), attaindrait à une meilleure compréhension. Dans ce cas toutefois, on peut se demander si la traduction conserverait encore une utilité, le lecteur possédant déjà l'information et/ou étant en mesure de prendre connaissance du texte original.

2. The French translation considered as an entry in a competition to recruit French translators

(1) Le texte comporte 7 ou 8 non-sens, c'est-à-dire des passages qui n'ont aucune signification.

En plus, il comporte 2 contre-sens, 2 faux-sens et au moins 17 termes impropres.

(2) Il semble que, de par sa nature, la traduction automatique ne soit pas en mesure de respecter ce que, selon la terminologie de la linguistique américaine, on appelle la "collocation", c'est-à-dire l'emploi de termes corrects dans une liaison verbe-substantif.

Ainsi, par exemple, on ne peut pas dire "les poules ont été divisées entre deux centres", mais l'on doit dire: "les poules ont été <u>réparties</u> entre ... etc".

(3) Dans une suite de plusieurs substantifs en anglais, la traduction automatique ne permet pas d'établir le rapport correct.

Ainsi le membre de phrase "the EEC Commission services" a été traduit : "les services de la CEE de la Commission" au lieu de "les services de la Commission de la CEE". Un traducteur n'aurait jamais pu commettre une telle erreur. Il en va de même jusqu'au cas où il n'y a que deux substantifs. Ainsi, "Poultry Breeding" a été traduit "volaille élevant" au lieu de "élevage de la volaille".

(4)   Plusieurs termes n'ont pas été traduits parce que, probablement, ils ne figuraient pas dans la mémoire.

Ainsi, "stunner", "bleeded", "hand-plucked" etc... (avant-dernier paragraphe de la page 2).

(5)   Plusieurs choses sont complètement incompréhensibles.
     Ainsi, (3ème alinéa de la page 1) "sabot", "camion
     anglais", "meter" traduit par "compteur" (3ème alinéa de
     la page 2) etc... ou encore "Roman I" laissé tel quel.

Conclusion

     Si ,je devais apprécier cette traduction comme une épreuve
de concours, ,1e me serais probablement arrêté au milieu de la
première page en mettant un zéro au candidat.

     En tout cas, je me suis beaucoup amusé en lisant que
"les poulets soumis à l'épreuve étaient capables de griller
des gâteaux de haricots de soja" et que "le 25ème poulet
n'était qu'un figurant"("an extra"), tandis que les 24 autres
devaient être probablement des artistes, alors que c'est pré-
cisément la machine à traduire qui a manqué son numéro de
trapèze et s'est écrasée au sol.

5. The French text revised

     Première conclusion qui s'impose d'emblée: non seule-
ment cette traduction se situe bien en-dessous du seuil de
rentabilité, mais on peut même la considérer comme franche-
ment inutilisable.

     Cela dit, la machine en question n'était visiblement
pas prête à ce genre d'expérience.

(1)   Son vocabulaire comporte des erreurs matérielles
     grossières. Je pense aux fautes d'orthographe; ex. :
     assomer ("assommer), abatoir (abbatoir).

(2)   Elle manque d'informations :

     a) vocabulaire : elle ne connaît pas des termes
        élémentaires comme "bleeded", "hand-plucked",
        "deep-frozen".

     b) morphologie : ex. : "23rd novembre","étude pour
        les poulets", "déterminer de l'eau étrangère".

     c) syntaxe : ex. : "aider à la planification et en
        effectuant l'étude étaient également dr. P.
        Stevens"

(3) Il aurait fallu lui apprendre à respecter les en-
    sembles de mots en langue étrangère sans chercher à les
    interpréter comme s'il s'agissait de mots anglais.
    ex. : "van" qui est traduit par "camion" alors qu'il

s'agit tout simplement de la préposition "de"; "Hoof" qui est traduit par "sabot" alors qu'il s'agit d'un nom propre; "für", qui devient "fourrure" alors que, là aussi, il s'agit d'une simple préposition.

On aurait ainsi évité des assemblages de mots hétéroclites tels que "la fourrure Fleischforschung De Bundesanstalt".

Ce dernier cas illustre du reste le manque de rigueur de la machine car, optiquement, "für" est différent de "fur".

(4)   Ce qui m'amène à varier des bizarreries de la machine. Dans certains cas, la machine laisse un blanc. Ainsi, "deep-frozen" est traduit par "profond -". Dans d'autres, elle reproduit le mot tel quel; ex. : "bleeded", "hand-plucked". Pourquoi ?

(5)   Si on passe maintenant à un stade de difficulté supérieur, on peut dire :

> a) que le stock de synonymes dont dispose la machine est insuffisant. Il semblerait que, pour chaque terme, elle ne connaisse qu'un seul équivalent. Ainsi, "implementation" = "exécution", "planning" = "planification", "removal" = "élimination". Il est évident que, dans ces conditions, la machine est vouée à commettre des faux sens .

> b) Remarque accessoire : s'il fallait à tout prix ne retenir qu'un seul équivalent, encore fallait-il choisir le plus courant et donner pour "live" l'équivalent "vivant" (au lieu de "vif", réservé à certaines tournures). Pour "neck", il fallait, c'est évident, donner "cou" (et non "col", d'un usage plus limité).

> c) Pour pouvoir fournir une traduction valable, la machine devrait, non seulement disposer d'un stock de synonymes suffisant, mais aussi apprendre à les choisir en fonction du contexte, ce qui, pour un ordinateur incapable de prendre des raccourcis, suppose sans doute toute une séquence d'opérations complexes. Autrement dit, il lui faudrait une mémoire beaucoup plus développée.

(6)   Passons sur certains "gags" très réussis comme ce
"25th" poulet qui doit "être un figurant" ou alors cette
autre formule "tout avait été affamé pendant 12 heures
avant l'abattage", si drôle qu'elle confine à la poésie.

Il reste certains résultats inexplicables tels le
"camion anglais" dont, avec la meilleure volonté du mon-
de, on ne parvient pas à retrouver l'origine dans "Labo-
ratorium voor Hygiene en Technologie van Eetwaren van
dierlijke Oorsprong". De même, on comprend mal comment
"Station de Recherches Avicoles", en français dans l'ori-
ginal, a pu devenir "poste De Recherches Avicoles", après
avoir transité par la machine à traduire.

Conclusion

A mon avis, la machine qui a fait cette traduction
possède une mémoire beaucoup trop rudiment aire.

L'expérience est donc prématurée et ne peut pas être
concluante. Elle ne saurait l'être en ce sens qu'elle ne
permet pas d'apprécier les résultats que pourrait donner
une machine suffisamment équipée.

On peut le regretter d'autant plus qu'une telle ex-
périence risque d'apporter de l'eau au moulin des détrac-
teurs de la machine à traduire.

ANNEX II

<u>REPLY FORM FOR INTERROGATING "EURODICAUTOM"</u>

The form reproduced on the next page has been designed at the Council of the European Communities for noting the relevant portions of answers obtained during interrogation of "Eurodicautom" from a visual display terminal.

The full term requested is written out at the top of the form. If "Eurodicautom" contains it in full, as requested, a tick is placed in the "original" column in the first space provided for answers, and the translation is noted.

If the term requested does not appear as such, even after truncation, the relevant portions of subsequent answers are noted, in both "original" and "translation" columns. It will not normally be worth continuing further than the five answers for which space is provided on the form;

Since the system gives consecutive numbers to the answers it provides to any one question, and it may not be helpful to note them all down, the left-hand column on the form enables one to record the numbers of the answers which are noted, in case it is required to check them later.

The second column, headed BE ("bureau émetteur"), enables one to note the terminology bureau from which the answer originates, since "Eurodicautom" contains terms from several different sources and the source of each answer is indicated on the screen by code letters. Experience with using the system may in fact indicate that material from some terminology sources is more helpful or relevant than that from others, for one's particular purposes.

The heading of the final column is self-explanatory. It may be useful to note the context of the term provided in an answer, or some other information given on the screen, in order to help the translator make his choice from the information obtained.

* circle appropriate languages

Term requested	from E D F N I Da * to E D F N I Da *	Date	Time

ANSWERS

Doc No.	BE	Original	Translation	Notes