[From: *Overcoming the language barrier*, May 1977, vol.1 (München: Verlag Dokumentation, 1977)]

TERMINOLOGICAL ACTIVITIES IN THE EUROPEAN INSTITUTIONS, WITH SPECIAL REFERENCE TO EURODICAUTOM

J. GOETSCHALCKX
Head of the Terminology Bureau
Commission of the European Communities
Luxembourg

SUMMARY

For the European institutions, multilingualism is an inevitable daily reality. All the different language versions of decisions enacted are equally authentic and directly applicable in Member States.

Their translation services are large, and have not only to work in the six official languages, but also translate documents dealing with widely differing subject fields, at a very high level of technicality and specialisation,

For that reason, terminology and documentation services have been attached to the translation services.

The Heads of the translation and terminology services meet regularly at sessions of the GIIT ('Groupe Interinstitutionnel de Terminologie et de documentation') to exchange information and coordinate their efforts.

Realising that the traditional reference works are no longer adequate for the task, the Commission has set up a computerised terminology system, EURODICAUTOM, which makes it possible to collect all available information and to disseminate it in the form of cards or glossaries, and also permits interrogation of the data base in the form of batches of questions or individual questions. This very flexible and simple system can accept all terminological information, whatever its form. On the other hand, conversational interrogation of the data base at the terminal screen is backed by a highly sophisticated search strategy.

1. THE NEED FOR TRANSLATION

'What's in a name ?', say the English. But if one speaks of a European Community it must of necessity be a multilingual community. This was, moreover, the option chosen by its founders, who, from the earliest meetings leading up to the present Community, accepted the principle of equality of statue for all the national languages of Member States. This rule is still in force, even though, internally, quite a few working documents are translated only into two or three languages.

Prom the moment the Community enters into its legislative role, that is to say when it takes decisions or makes recommendations, the rule requiring translation into all of the languages is strictly applied.

But the choice was not based on considerations of a national or sentimental nature, or regard for picturesque local tradition. Ever since the French revolution, and in keeping with the principle of nationhood, it has been almost unthinkable that laws should be drawn up in a language other than that of the citizen who is supposed to respect and apply them. For it must not be forgotten that nationals of the nine Member States are at the same time, in a sense, 'citizens' of the Community.

Unlike international organisations of the UNO type, whose recommendations or decisions become enforceable only through legislation at the national level, Community laws are directly applicable to all natural and legal persons in the Member States.

The managers of agricultural cooperatives in Denmark must know exactly what are their rights and obligations within the framework of the agricultural policy. The clerk of a pension fund in Calabria must know the rules governing the calculation of rights acquired by

a worker who has been employed in various Community countries. The migrant worker himself must know what are his rights with regard to social security benefits. Company lawyers planning an industrial combination must be able to understand every nuance of the texts published by the Community services responsible for problems of competition. Recently, the doctor at the end of his career, bored to distraction in the mists of Scotland, has needed to know what prospects are open to him if he wants to warm his old bones in the sun of Italy or the south of France while still working as a doctor, even if only to round off the income from his savings.

2. ORGANISATION OF TRANSLATION SERVICES AND TERMINOLOGY

Thus, for the Community, translation is not a luxury but a necessity, a vital need. That is why each European institution has its own translation service, with a specialist group for each language : German, French, Italian, Dutch, English, and Danish. Every translator translates from several languages into his mother tongue.

Likewise with terminological activities, each of the Institutions does some work in this field, in line with its main or most pressing needs. Thus, the Council of Ministers has a Terminology Committee to coordinate the terminological work done at the level of the Language Divisions. The European Parliament has its own Terminology Bureau. The Commission has a Translation Service at Brussels and a Medium- and Long-Term translation service at Luxembourg, each with its own Terminology Bureau. The Economic and Social Committee and the European Investment Bank have translation services, but terminology, although practiced there as in the other Community services, does not yet enjoy official recognition.

Although every terminologist must conform to the needs of the translation service to which he belongs, it is clear that these many activities converge in some degree, and thus require coordination. This is the role of the GIIT, the Groupe Interinstitutionnel de terminologie et de documentation, which brings together representatives from the translation and terminology services of all Institutions of the European Community.

3. LEVELS OF TRANSLATION

If needs differ from one institution to another they also differ from one document to another, for there are various levels of translation.

3.1. First, there is the whole range of day-to-day communications, along the style of the 'Staff Circular', which are seldom urgent and pose no special problems, particularly from the terminological standpoint.

3.2. At the other end of the scale, as far as difficulty of terminology is concerned, are the reports and similar documents presented by national experts to working parties of the Communities' committees. These texts are very specialised, and often of a highly scientific or technical nature. The same holds for the reports presented at scientific meetings and at colloquis and congresses. Here we are faced with primary information, and all the terminology work connected with translating the documents, so that they can be discussed at various levels, still remains to be done.

3.3. Another type of text is the documentary text. These come in the form of a series of keywords, or very concise texts of high information content. In its various capacities, the Community is concerned with a wide range of documentation systems. Here again, the problems of terminology have received no attention up to the time when the documents are sent for translation.

3.4. Examples of information texts intended for a wider public are the general report, the report on social activities, monthly bulletins, etc., etc. Generally, they do not cover new ground. The terminology needs checking only to ensure uniformity.

3.5. Texts intended to mould opinion occupy a special place. Addressed to governments, the general public, or certain economic or social categories, they raise few problems of terminology, but make greater demands on the full resources of the language, both vocabulary and syntax.

3.6. The democratic nature of the Community requires that economic sectors and social groups, as well as the nations as such, shall have their say in all measures envisaged by the institutional apparatus.

The minutes of meetings of the Economic and Social Committee, the Consultative Committee, or the committees of Parliament, and the parliamentary debates themselves amount to a large number of pages to be translated.

Here again there are few problems of terminology, although the oral nature of the texts may lead to difficulty in translation. On the other hand, they often refer back to earlier proceedings and to various bodies within or connected with the Communities. This work is therefore more in the nature of documentation. Nevertheless, it is frequently taken on by the terminology service. With texts of this type, and above all the parliamentary debates, the great difficulty lies in the very close deadlines for translation.

3.7. All these activities result sooner or later in decisions, resolutions, or recommendations which, as we have already seen, have direct force of law in the Member States. Apart from the very strict rules governing the form of texts of this kind, it is essential to achieve absolute equivalence in the versions in the various languages, because all versions are equally authentic. In translating these texts, translators and revisers take on a heavy responsibility.

4. PROBLEMS OF TERMINOLOGY

It is clear from what has been said above that problems of terminology differ considerably according to the category of the document to be translated. Other problems may stem from the particular situation of the translator working in a large international organisation.

4.1. Difficulties due to the nature of the documents

4.1.1. Documents of a legal nature

In the case of documents of a legislative nature, the translator must be able to find the technical terminology which is already in the preparatory documents, and must adhere to the general legal-administrative terminology when couching all this in the form of a decision or recommendation. Of course, it is important that the titles or official names of earlier provisions, and of the bodies cited in the text, are left unchanged.

4.1.2. Scientific and technical documents

In the case of documents of a scientific and technical nature it is necessary to find out what is the correct terminology, and sometimes even to create it. This happens especially when concepts are to be translated into a language which has never previously encountered them. There is the well known case of mining terminology, which is practically non-existent in Danish. The same applies in the more general case where research reports or papers on recent progress in a science or technique are to be translated for congresses and other meetings of that sort. But the same situation can arise in a more acute and difficult form in the legal field, because legal concepts have often evolved along very different lines under the constraints imposed by political, philosophical, sociological, and religious doctrines, and even by geographical conditions.

4.1.3. Economic and political documents

Economic or political documents produced by the services of the Communities are often the point of departure or the result of

delicate negotiations and discussions. In these documents it is often important to ensure a certain uniformity. Those who are familiar with the mysterious alchemy of the Communities tend to believe that each variation in a statement corresponds to a different reality or at least to an inflection of that reality. By inadvertently changing a term one runs the risk of laying the entire matter open for rediscussion.

In addition, lack of uniformity may cause difficulties in indexing. There are difficulties enough with the monthly bulletins, where the original text is often in French. If the translations into the other languages are not uniform, a concept always expressed by the same word in French may occur in several different forms in one or several of the other languages. The user needs a fertile imagination to know how the concept will appear in 'his' index. Clearly, cross-indexing can solve this problem, but it is nonetheless true that uniformity of terminology will promote clarity and simplicity, and hence simplification.

4.2. Difficulties due to the translator's situation

Several situations are possible for the translator :

1. he understands the original text, but has difficulty in translating it into the target language ;
2. he does not understand the original text ;
3. he is translating from a language of which he has only an imperfect knowledge.

In the first case the difficulty may reside in the use of the correct terminology. It is not possible to know the terminology of all fields. It may also be more a question of the syntax used by the specialist in a given field. In many cases the question of how to use the appropriate verb or preposition gives more difficulty than the choice of the correct term. Should one say 'appareil à pression', 'appareil sous pression', or 'appareil de pression' ?

If the translator does not understand the text, and above all if the document relates to a matter with which he is not familiar, things become complicated. This happens frequently because many

subjects are covered, and the volume of documents to be translated can vary very appreciably from day to day. One day there may be mining documents for half the team, while another day it is 'medicine' for everybody. For each field there can only be one or two specialised translators in the team, and then their combination of languages must be the right one.

The third case may also crop up when the volume of documents to translate from a given language is beyond the capacity of those having a perfect knowledge of that language. Then it becomes necessary to turn to people for whom it may be their third, fourth, or even fifth language.

Speaking generally, terminology as applied to translation must, thus, either ensure uniformity,
or ensure authenticity,
or ensure conformity with the usage of the professional.

Having stated the problem, let us now see how translators, revisers, and terminologists attempt to solve it.

5. POSSIBLE SOLUTIONS FOR PROBLEMS OF TRANSLATION

Generally, translators and revisers look to their own sources of information before enlisting the aid of terminologists.

They call on the service offered by a bureau of terminology or documentation only when their own knowledge and experience, and the dictionaries, documents and reference works at their disposal, and also colleagues and the experts they know personally, have failed to come up with the answer.

5.1. Documentation

In the face of these difficulties the first step is to try to collect useful reference documents.

The individual translator, the section or the division, and the terminology bureau, each according to his or its needs and means, collect documentation comprising the Community documents which constitute the basis of any future action.

Documents published by other international organisations also constitute a useful source of information. It may be that a matter which is now starting to interest the Community has already been studied, in whole or in part, by other organisations such as UNO, OECD, or the Council of Europe, to mention only three.

Reference works including translation dictionaries, single language dictionaries, reference works of a fairly general nature, technical manuals, research reports, etc., etc. are available at the three levels mentioned above.

The rapid evolution of the modern world is a severe challenge for the publisher : a book which appears on the market was completed at least two years earlier, and much has happened in the mean time. This is why the abstracting of technical journals is a very important link in the chain of documentation set up for the benefit of the translator. Card indexes arranged by keywords or a systematic classification provide access to the latest information in the journals.

5.2. Terminology card indexes

5.2.1. Usefulness

Ready-to-use terminology represents the most direct and effective aid. For the translator and the reviser the ideal is to find a card giving the term or expression for translation together with its equivalent in the target language.

5.2.2. Setting up a card index

The indexes are based on research done by terminologists in response to questions put to them, and also on information gathered by them when scanning technical reference works and articles, specialist journals and reviews, etc.

5.2.3. The terminology card

Opinions differ greatly as to the minimum requirements which a terminology card should satisfy. These opinions vary according

to the needs to be met, the 'customers' who use the system, and the means at the disposal of the service which maintains the index.

5.2.3.1. Terminological information

- term-to-term card : this gives an expression or word with its equivalent in two or more languages. The expression may be a Uniterm or a multiterm,
- card with definition or illustrative context : this gives a term accompanied by a definition or by an explanation of its meaning and use. It may also include one or more illustrative contexts, showing several uses of the vedette,
- phraseological card : this gives equivalent sentences or parts of sentences in two or more languages.

5.2.3.2. Information on documentation

5.2.3.2.1. The card should indicate the source of the information given. This makes it possible to check back if, for example, the information is challenged. It can also give the translator a useful lead to information which may help him solve other problems of terminology in the same document.

5.2.3.2.2. The card should have one or more subject codes indicating the field(s) in which the term or expression is used.

5.2.3.2.3. It can be very useful to add a note giving other information of interest : etymology, references to related terms, statement of the geographical or political area in which the term is used, etc.

5.2.4. The justification for different types of card entries

5.2.4.1. The term-to-term card

The term-to-term approach is adopted in many, if not most glossaries, dictionaries, and other similar works. It is also the kind of card which the translator, working under pressure, often makes for his own use.

To refuse this kind of information is to abstain from access to a large amount of terminology which is often perfectly valid.

Equally clearly, using a card of this sort entails risks which cannot be ignored if it comes into the hands of a translator who is not familiar with the subject.

On the other hand, it is acceptable when one is dealing with an absolutely unequivocal term such as fluorescence spectrography.

The risks entailed with a card of this sort can be very appreciably reduced by adding subject codes, i.e. codes indicating the field(s) in which the term is used.

All the same, polysemy is always lying in wait. An English card giving the mining term 'roof' with 'toit, couronne' as the French equivalent would be useless for a French translator unfamiliar with mining terminology. He could not know that 'roof' is translated as 'toit' in the working area, but by 'couronne' in a roadway. Here, the subject code would not help him unless it was made so specific that it would no longer be possible to learn it by heart.

5.2.4.2. Card with definition or illustrative context

This kind of card is extremely useful for indicating the difference between closely similar concepts, which often confuse the uninitiated, and, as we have seen, the translator is often in this category.

On the other hand, the definitions are often questioned. Each expert considers his own definition to be the only valid one, and

it often happens that in changing from one language area to another one is confronted with definitions covering different semantic fields.

To avoid surprises of this sort, it is necessary to use the definitions supplied by standards institutions. Unfortunately, the work of standardisation progresses extremely slowly. Often, concepts used regularly in everyday language have not yet been defined. Certain ISO working parties take years to define a hundred or two hundred terms. The infrequency of the meetings and the wide geographic scatter of the experts makes it impossible to do better than this.

Those who know ISO well also know what difficulty there sometimes is to arrive at general agreement on the wording of a definition. That, by the way, is the reason why quite a few of the definitions are almost incomprehensible. A good example of this is the definition of 'minière' in French mining law: "Les minières comprennent, d'une part, les tourbières et, d'autre part, les gîtes de minerai de fer autres que ceux dont l'exploitation à ciel ouvert cesse d'être possible sans l'etablissement de puits, galeries et travaux d'art ou dont cette exploitation, quoique possible encore, doit durer peu d'années et rendre ensuite impossible l'exploitation avec puits et galeries."

In cases where terminological research is concerned not solely with the semantic aspect but also with the syntactic aspect, the definition alone is not enough.

An illustrative context is very useful when giving examples of the use of terms presented in the vedette form. However, as these are authentic contexts, there is no concordance or equivalence, with the result that the translator can only obtain information of a syntactic nature.

5.2.4.3. The phraseological card

More and more, theoreticians and practitioners are turning towards phraseology.

Linguists are increasingly of the opinion that true equivalence is to be found at the level of the sentence. 'The relation of equivalence holds between a sentence in one language and a sentence in another language ... if each of them is an optimal translation of the other in a given context' (Marton).

In scientific and technical documents the difficulties lie not only in the use of a terminology which is known only to the initiated, but also in certain syntactic characteristics or the fact that the professional gives the words of the language a special meaning, particular to his field.

From what has been said it emerges clearly that there are in reality specialist languages, and it is therefore necessary to juxtapose sentences and parts of sentences in order to obtain equivalent statements which are at the same time in conformity with professional usage in the languages in question. For the translator cannot be content simply to reproduce a restructured or unchanged sequence of words having the same meaning. His translation must be presented in a form 'that accounts for the functioning of language in man' (Sciarone - ITL 32/9).

The translator is well aware of this. The French translator, for example, when checking terminology, will first of all read the original text and then ask himself 'how do we say that in French?'. On the other hand he will reject a not particularly felicitous formulation with the argument 'we don't say that'. In the technical field, therefore, one must be specific and ask 'how does a miner, an iron and steel specialist, a doctor, or a lawyer, etc. say ...?'

Moreover, terminological information presented in the form of equivalent sentences, being richer in content than that in the word-to-word form, has the extra advantage that it often conveys scientific or technical information on the subject in question. As wo have seen that translators only rarely have the opportunity to specialise in several fields, any information which gives them a better understanding of the problems in the field covered by the document to be translated cannot fail to help them provide a

translation which keeps to the language of the professional, the ideal always being a translation which reads like an original text.

Moreover, the phraseological approach has the advantage that it may sometimes provide information of which the translator believed himself in no need, although with this attitude he was in danger of being unfaithful to his duty as a translator, that is to say a linguist of undoubted experience, but not a specialist in the matter in question. In a way, this would be the opposite of the old saying 'traduttore ... traditore'.

Let us take a look at the following example:
In fretting corrosion the surface of the material is continuously polished by abrasion.
Dans le cas de corrosion par frottement, la surface du matériau est continuellement avivée à cause de frottement.

The translator who has the task of translating this English sentence into French will probably look for the French equivalent of the expression 'fretting corrosion', but it is unlikely that he will check on the translation of the verb 'to polish'. But it is equally unlikely that he would have thought of translating 'polish' by 'aviver' unless he had special knowledge of the field.

Nevertheless, using 'aviver' will bring him closer to his ideal, a translation which reads like an original.

By way of conclusion, one may say without any doubt that card indexes are very useful working tools and even indispensable to the translator. However, they have the disadvantage that it takes quite a lot of work to keep them up to date, and duplication is practically impossible.

Of course, people have their individual card indexes, but these are not generally accessible to colleagues.

6. GLOSSARIES

Glossaries offer the same type of information as cards, but all on a given subject: treaties, budgets, maritime law, patents, atmospheric pollution, steel in the building industry, etc. For that reason the terms do not always need to be accompanied by a subject code, as they are already defined by the overall theme.

Among glossaries, a distinction must be drawn between citation glossaries, which is to say those based on documents which exist in several language versions, and original glossaries, which is to say those established on the basis of original texts by juxtaposing corresponding or equivalent terms or contexts.

In the first-mentioned, the terminological research has already been done beforehand, as the different language versions already exist. In the second case it is necessary to search for documentation in all the languages, and then analyse it and establish concordances.

The object of citation glossaries is to ensure the authenticity and uniformity of documents produced at a later date.

Original glossaries are intended to solve terminological problems which arise when it is necessary to translate documents dealing with matters which have not yet been the subject of action by the Community, or which deal with very recent progress, or even cover entirely new ground.

In the twenty years of their existence, the terminology services of the Community institutions have produced about 100 glossaries.

Although in principle all glossaries deal with different subject fields, the translator confronted with a given terminology problem is not always clear as to which of the available glossaries is the one in which he has a chance of finding the answer. Moreover, if he has consulted a glossary in vain two or three times he will reject it in this entirety for ever, thus permanently depriving himself of all the useful information contained in it.

7. FIRST ATTEMPTS AT AUTOMATION

7.1. DICAUTOM

Mr. J.A. BACHRACH, the first Head of the Terminology Bureau, of the High Authority of the E.C.S.C., and at present Head of the Medium- and Long-Term Translation Service at Luxembourg, was already aware of these problems when he assumed responsibility for that bureau in 1963.

That was why the DICAUTOM system, which was the result of collaboration with the Centre de linguistique automatique appliquée of the Free University of Brussels, headed by Mme Lydia HIRSCHBERG, was conceived from the outset as an automatic dictionary to take the place of all the glossaries, dictionaries, and other working documents.

The system evolved and was developed progressively until in 1968 it emerged as an operational system capable of accepting questions couched in phraseological form in German. Rules of morphological reduction made it possible to reduce each word of the question to its standard form. Then the system searched among its stock of sample sentences for those corresponding most closely to the question. Analogous systems of morphological reduction were already being developed for French, Dutch and Italian.

7.2. EUROTERM

At the moment the High Authority of the E.C.S.C. in Luxembourg gave the go-ahead to the DICAUTOM system, the Commission of the Economic Community, often called the Common Market, gave Mr. F. CUYPERS in Brussels the opportunity to launch the EUROTERM system, which was in effect a presentation of fundamental texts — treaties and important regulations - in the form of KWIC lists. Because of the high output by the Community services of texts of this sort, EUROTERM very rapidly became an extremely unwieldy instrument, impossible to handle, at least in the form of computer listings.

7.3. EURATOM glossary

The Commission de l'énergie atomique, better known under the acronym EURATOM, which was charged, under the terms of the treaty which established it, with the task of deciding nuclear terminology, had invited its Terminology Bureau, headed by Mr. H. KOWAISKI, now Chief of the combined Terminology Bureau at Brussels, to produce the EURATOM glossary.

This glossary can hardly be considered as a form of automated terminology, as it comprised no software for the treatment of terminological information. The computer was used only for sorting and printing.

Unfortunately, the merger of the Executives and the change of computer put a stop to all this work in 1968.

This most regrettable interruption did not prevent the continuation of the work. The Terminology Bureau of the Commission at Luxembourg continued to prepare glossaries conceived for computer storage sooner or later, and to study all problems relating to the use of computer for automated terminology.

8. EURODICAUTOM

8.1. History

When, in July 1969, the Terminology Bureau at Luxembourg was instructed by the Director General of Personnel and Administration, Mr. Lamberto LAMBERT, to take on all work relating to the automation of terminology, and to combine all existing systems in a system to be known as EURODICAUTOM, it had beyond any doubt the necessary knowledge for taking on the task, but did not have the indispensable practical means.

Finally, in 1973, the Management Committee of the Computer Centre gave the green light for collaboration with Directorate General XIII to launch the new EURODICAUTOM system.

The umbilical cord between DICAUTOM and EURODICAUTOM consists in the shared principles which were the basis of their conception from the point of view of terminology. At the level of data processing, the approach is very different, which is quite normal because in

the mean time technology has advanced.

The EURODICAUTOM system was based on close cooperation between the Terminology Bureau of the Commission at Luxembourg and the Analysis and Programming Unit of Directorate General XIII, headed by Mr. PIETTE, with Mr.FRANCOIS responsible for the EURODICAUTOM project.

As to its conception from the terminological point of view, there were regular contacts and follow-up between all interested services through GIIT, the Groupe Interinstitutionnel de Terminologie et de Documentation, and its specialised sub-groups 'Harmonisation' and 'Automation'. Because of this, EURODICAUTOM takes into account the needs of all the translation services of the Community.

8.2. Aims of the system

The system tends to pool all existing terminological information, whatever the form in which it is presented. On the other hand, the collection of information is decentralised, so that each collecting institution can feed the system while keeping its own identity. This also appreciably reduces the administrative work.

The system must be dynamic in the sense that it can be extended to other languages without great difficulty.

It must be simple, so that the user can make use of it without the aid of a specialist and without a long and complicated apprenticeship.

8.3. The EURODICAUTOM terminology document

8.3.1. Terminological and documentary content

Starting from the principle that the system must centralise all available terminological information, whatever the form in which it is presented, EURODICAUTOM accepts the three types of information already described: term-to-term, definition, and phraseology.

In addition, it can furnish, for each terminological entity, source references, subject codes, and, where appropriate, an explanatory note.

8.3.2. In Annex I you will see the layout of a terminology document, with an explanation of the codes used.

I must draw your attention to three codes which play a special and very important role. These are: the code BE ('Bureau émetteur'), indicating which Bureau originated or collected the item of information.

The code TY ('Type') indicates whether the item of information in question is an individual one ('fiche isolée' = FI) or whether it comes from a glossary or some special file, for example.

NI ('numéro d'identification') is the identity number, that is to say simply a serial number.

Given that each terminological entity is defined by its BE-TY-NI, that is to say the combination of the three elements mentioned above, the collection of items of information can be decentralised, because the numbering of the entities is independent for each originating bureau.

Thus, the insertion of a particular collection of items of terminological information, such as a glossary or some special file, can be done with independent numbering, so that the collection retains its autonomy within the system. In effect, the authorisation of a distinctive TY code avoids all confusion. This element also plays a role in searches, as we shall see later.

8.4. Data input

Data input follows the form which you will find in Annex I, already mentioned.

The digitisation of the information is done by optical character reading. Punched card input was not retained because the punching capacity of the Computer Centre is inadequate. Because the input is in many languages, the work could not be done outside the Commission. To have each language punched in the corresponding country would cause administrative problems incompatible with the smooth running of the system as a whole.

The solution adopted has the advantage of being relatively cheap, as the typing of the input forms can be done on the spot, directly from the documents prepared by the terminologists. In this way we save ourselves the cost of typing a clean copy beforehand. Also, in case of difficulty, the secretaries can go straight to the terminologist.

For certain tasks of a rather special nature, input is by flexo-writer.

New input equipment of the encoder type will be in use in the near future.

Some data obtained by exchange with other automated systems is fed in directly by magnetic tape.

8.5. Corrections

Correction can be done at two stages:

a) after preparation of the sheets for optical character reading

b) on the computer printouts furnished by the Computer Centre after optical character reading.

Correction at the first stage is sometimes desirable when the typing has been done from manuscripts.
Correction at the second stage is indispensable.

After correction of the printouts, the necessary changes are entered on an 80—column sheet. The actual correction is done by punched cards. The 80-column sheets first identify the incorrect text, and then show it in its corrected form.

For correction, the code NI is replaced by the code NX.

8.6. Subject coding

Subject classification is indicated by subject codes. The subject codes accompanying each entity are those of the Lenoch classification. The subject classification developed by Dr Lenoch, who has for many years been a member of the staff of the Terminology Bureau of the Commission at Luxembourg, is based on the UDC classification, but is composed of mnemonic codes based on the French wording of the main headings. Take for example 'steel bridge':
code BA9 stands for bridge
code BAA indicates a connection with roads
code SIJ indicates that it is made of metal.

Thanks to these codes it is possible to know the field of use of a terminological entity, even if it is presented as an isolated term. There is no need to consult a list or table of codes. The 45 main headings are easy to learn by heart. Nevertheless, those who use the system only rarely, and so lack the practice to be able to recognize all the codes immediately, can interrogate the system to find the meaning of the codes which they have seen on the screen in the CM ('code matière') zone.

For entities presented in the phraseological form, or as a vedette accompanied by a definition, there will of course be no problem of this sort, which is an indication of the superiority of these two approaches as compared with term-to-term information.

These codes make it possible to produce glossaries on particular subject fields, as we shall see later.

8.7. The content of the data base

Up to the present time, the EURODICAUTOM system contains only the glossaries produced by the Terminology Bureau at Luxembourg (BTL) and two glossaries produced by the Brussels Bureau (BTB), namely those covering the Treaties and European Standards.

Brussels are beginning to feed in their glossaries. The Parliament intends to produce a new glossary, using the EURODICAUTOM system. The Council is already supplying terminology.

Many other organisations have given us permission to feed their terminology into the system.

The first to help us was the Terminology Data Bank of the University of Montreal, which placed some 70 000 entities at our disposal.

It must not be forgotten that every day the translators of the Community Institutions do a considerable amount of terminology work which, as matters now stand, is lost as far as their colleagues as a whole are concerned.

Use of the standard card shown in Annex II would enable all translators to benefit from the work done by each of them, without imposing extra work.

As items of terminological information differ in value, a reliability code ('code de fiabilité') is provided, on a scale running from 0 to 5. Code 5 indicates terminology which is mandatory, such as standards or quotations from Treaties, regulations, and other texts of that sort.

Code 4 indicates a complete card, i.e. a card having all the necessary elements : contexts or definitions and sources.

8.8. Elimination of redundancy

When terminological information from very different sources is fed into the system, there is a certain risk of redundancy.

Sometimes this redundancy is only partial. For a particular concept there may, for example, be a Dutch/Italian document and a French/English. One terminology document may contain a term standing alone, while another contains the same term, but with a definition. But perhaps the first one contained an additional language, etc., etc. In this case it is obviously necessary to combine the two documents.

The operation of eliminating or combining redundant documents is carried out by the terminologist, working from a computer printout. The printout is obtained by a type of serial interrogation in which each entity is treated as a question. All documents corresponding in whole or in part to the entity in question are thus brought together and can easily be examined so that their fate may be decided.

8.9. Various possible types of output

8.9.1. Printouts (lists)

Terminology documents stored in the computer can be retrieved in the form of printouts. These are used in the main for correction and completion work, for example the supply of missing languages.

8.9.2. There is also a programme for the production of cards. This makes it possible to exchange information with organisations having no computer.

8.9.3. Glossaries

When particular subject fields become topical, and there is an appreciable rise in the volume of documents for translation from those fields, it can be useful to supply the translators with appropriate glossaries.

The same applies to translators and interpreters who have to be present at congresses outside the European institutions, so that they are unable to consult the terminology data bank. A glossary on the subject field of the congress is a very valuable aid in these cases. By use of the subject codes it is an easy matter to produce a glossary comprising text and alphabetical index.

8.9.4. Magnetic tapes

Magnetic types are the basis of exchanges with other terminology data banks.

8.9.5. Batch interrogation

In the case of non-urgent translations, the terminologist can put one or several batches of questions to the computer. The answers are printed out on the following day on the rapid printer at the Computer Centre, and delivered by internal post.

8.9.6. Interrogation at the terminal screen

8.9-6.1 Search strategy

For urgent translations and for specific questions, conversational interrogation is best.

As the system was conceived by translators for translators, the accent is on simplicity and speed in use.

The simplicity of the system is such that an hour's training at the terminal screen is enough.

When the user, by entering the code TERM, has indicated his wish to consult the terminology data base, and has then entered his personal password and surname, the instructions which appear on the screen guide him through all subsequent operations.

The procedure is very simple :

1. indicate source language,
2. indicate target language(s). If you want all languages, enter ALL,
3. put question.

The question is put in the form in which it occurs in the text.

If the exact answer is present in the data bank it will appear on the screen almost instantaneously.

If the system contains only a part of the answer, it will produce what it has. Our philosophy is that in some cases a part answer can be useful to the translator. As the system is programmed to give the best answer first, followed by others in descending order to pertinence, the translator realises at a certain moment that the information being supplied is no longer useful, and then decides not to continue with the question.

The progressive decline in the pertinence of the successive answers supplied is a deliberate feature of the system, and is not reversible.

The reason for this is that the sequence of the answers supplied is determined by a system of weighting which takes into account the nature of the available terminological information corresponding to the question which has been put, for example whether a term is in the form of a vedette, or a keyword in context, or whatever. The origin of the information is also taken into account. Clearly, information originating within the Communities is more likely to meet our translator's needs than information from other sources. The target language requested is another determining factor. Further parameters are the subject code, the reliability code, and the type (TY).

8.9.6.2. Truncation

Since questions are put in the form in which they occur in the documents, a translator can put the question 'aciers inoxydables' because that expression is often used in the plural form. However, if the term is present in the data bank in the singular, there will be no answer, although in fact the elements of a useful answer are present.

The use of truncation, which allows the removal of a certain number of letters from the end of a question word and their replacement by others, will thus produce the answer in the singular although the question was in the plural.

This facility is particularly useful in the case of inflected languages such as German, and also in a language like Danish, in which the definite article may be postposed.

Up to the present time, truncation has been effected in a more or less mechanical fashion, mainly depending on the length of the word. We intend at a future date to introduce a system of truncation based on criteria appropriate to each particular language.

When interrogating the data base, the user can obtain truncation simply by pressing a button.

8.10. Advantages of the terminology data bank

Like the most beautiful girl in the world, a terminology data base can only give what she has. The same can be said of dictionaries, glossaries, and card indexes. However, the data base does have numerous advantages :

8.10.1. Rapidity

We should not compare a data base with a cupboard full of dictionaries and glossaries which the translator must consult one by one in order to find the solution to his problem. In that comparison the advantage of the data base is too obvious.

Let us think instead of the translator faced with a large card index, and the time it will take him to find the drawer which should contain the word he is looking for, then open the drawer, then search through the cards. In that time he could easily have put three questions to the terminal.

3.10.2. Centralisation

The storage capacity of a computer makes it possible to stock a large amount of information while ensuring easy and rapid access.

Up to now, each translator and each terminology bureau has had an individual card index. To take such an index and supply copies of it to everybody would be quite impossible, but feeding the indexes into a computer would make it possible to arrive at the same result with much less effort.

3.10.3. Exchange facilities

As input is the major problem with all terminology data bases,

automation has the enormous advantage that exchange of information by magnetic tape can be done without great difficulty and is not expensive.

8.10.4. Publications

With automation, glossaries and collections of cards on a particular subject can be produced quickly and easily, especially in the case of certain political and scientific meetings which take place outside the institutions of the Communities.

8.10.5. Facilities for updating

Unlike traditional card systems, where there are, inevitably, main and subsidiary cards, a computerised system has a single document for a given terminological entity, so that any changes or additions need to be made only once.

9. CONCLUSIONS

As long as the translator can solve his own terminological problems, either by personal contacts or by reference to the reference works, dictionaries, etc. at his disposal, there are no problems for the terminologist because he is not needed.

Thus, the terminology data base in no way diminishes the intellectual role of the translator. Nor does it mean an end to dictionaries, at least until every translator has his own terminal, and we are still a long way from that.

On the other hand, it does give us a chance to bring together for everybody's benefit the terminology collected by individuals. This collaboration removes not only the partitions which separate translators, but also the geographical and other barriers between the various European institutions and all other organisations which are trying to equip themselves adequately to meet the challenge of multilingualism.

Thus, the translator will be able to give his entire energy to the task of translation without having to trouble himself with terminology problems such as those which arose at Babel when the workers gave different names to the bricks and building equipment. And yet, 'what's in a name?'.

ANNEX I

EXPLANATORY NOTES ON THE OPTICALLY SCANNED INPUT FORM FOR TERMINOLOGICAL ENTITIES IN THE EURODICAUTOM SYSTEM

NI (NX) The number indicating the numerical sequence of the entity and permitting its retrieval. The code NX together with the number of an item previously entered with the code NI, allows corrections to be made to the original item or further material added to it.

TY = type Indicates whether the item is contained in a glossary (e.g. RPT = rayons parasites Toulouse, referring to the Conference on Parasitic Radiation held at Toulouse in 1969) or a single card containing the results of terminological research carried out in response to specific questions. E.g. TFI 0237

BE = bureau émetteur (bureau of origin) Indicates the bureau which produced the terminological entity. This helps the questioner to make a choice between cards giving different information. For example, a translator may have more confidence in a card on a financial topic from the EIB or IMP than from other sources. This indication also allows the user to bear possible national or local peculiarities in mind (e.g. possible differences in French terminology in Canada, France, Belgium or Luxembourg).

In addition, when the terminologists wish to produce a glossary on a given subject they must know where the information given by the computer comes from.

CF = code de fiabilité (realiability code) Indicates the reliability of the information given - not from the point of view of its intrinsic quality, since the experts themselves often disagree, but according to the presence of absence of valid references. For the time being, it serves to put the user on his guard and to help us improve information contained in the computer by adding or eliminating terms, as appropriate. This we are beginning to do.

The code 0 is given to an item for which no source is available, and the code 5 to information taken from standards or equivalent documents, e.g. treaties of the European Communities, which have a mandatory character. The intermediate codes indicate that there is no source of information in some languages.

AU = author The name of the author of a terminological card. This is used in cases where translators, interpreters or other persons assisting us on an occasional and voluntary basis wish to make a contribution and would like their contribution to be identified in this manner.

CM = code Matière (subject code) The subject codes are taken from a system of classification devised by a member of the Terminology Bureau in Luxembourg, Dr Lenoch. It is a three-position alpha-numerical code, usually of a mnemonic nature, based on the Universal Decimal Classification and continually adapted in accordance with our own needs and the requirements of the computer.

Examples: AD4 = administrative staff
MES = nuclear medicine

The purpose of these codes is to make it easier to pin down the field required during an on-line interrogation session, and they are particularly useful when producing subject glossaries.

PH = phrase Indicates a phraseological entry or illustrative context.

DF = definition There may be one or more definitions together with one or more contexts.

VE = vedette An expression of varying complexity constituting a "technical term" in the broadest sense.

NT - note Permits the inclusion of any necessary explanations or details. E.g. unusual plural, regional usage, obsolete term, etymology, etc.

RF = references Indicates the source from which the terminological information was taken.

It is essential to include the sources since this enables the user to verify the information if it is challenged an gives the translator with average experience an indication of the real value of the information offered. Finally, it also provides a documentation system, since if it is found from consulting the cards that several answers were taken from a given source, it would be reasonable to expect that answers to the as yet unanswered questions on the same subject might be found by returning to this source.

ENTITÉ TERMINOLOGIQUE

<table>
<tr><td>% NI (NX)</td><td colspan="3"></td><td>% TY</td><td></td><td>% BE</td><td></td></tr>
<tr><td>% CF</td><td></td><td>% AN</td><td></td><td>% AU</td><td colspan="3"></td></tr>
<tr><td>% CM</td><td colspan="7"></td></tr>
<tr><td>% IT</td><td>% PH</td><td>% DF</td><td colspan="5"></td></tr>
<tr><td colspan="8"></td></tr>
<tr><td>% VE</td><td>% MC</td><td colspan="6"></td></tr>
<tr><td colspan="8"></td></tr>
<tr><td>% NT</td><td colspan="7"></td></tr>
<tr><td colspan="8"></td></tr>
<tr><td>% RF</td><td colspan="7"></td></tr>
</table>