Terminology Resources in the Context of a Major Translation Project

Maria Rzewuska

Office of the Committee for European Integration
Al. Ujazdowskie 9; 00-918 Warsaw, Poland
e-mail: maria_rzewuska@mail.ukie.gov.pl

Abstract

The project consisting in the translation into Polish of Community legislation is run by the Office of the Committee for European Integration. It concerns ca. 60 thousand pages (as published in the Official Journal of the European Communities) and it is due to end by mid 2003. The quality of legal translations is the most important factor and a particular approach is adopted in order to ensure terminology consistency and legal correctness of these texts. The project comprises several stages – translation of primary legislation followed by the translation of secondary legislation and finally the translation of case law. Having in mind how important terminology consistency is, not only within one text but also within the whole project, the necessity for the elaboration of common terminology resources became clear. The project team has as its disposal a terminology bank and various specialised thematic collections of terms. The main repository of terms comprises general European vocabulary and is accessible to all free-lance translators.

1. Introduction

Poland struggling for membership in the European Union bound itself to implement Community law. This implementation of Community provisions and other measures undertaken in order to prepare Poland for its membership in the EU, have indicated the necessity to translate EU legislation into Polish.

Planning to join the policy of active multilingualism of the European Communities (McCluskey 2001) clearly indicated the need to dispose of EU legislation in new languages. Although several Candidate Countries were concerned by the same problem, they were left free to adopt their own solution leading to one and the same goal – translating EU legal acts into their official languages.

In the case of Poland a translation coordination unit was set up in spring 1997 at the Office of the Committee for European Integration. The team gradually expanded to count at present 20 persons within the frame of the Translation Department.

The overall translation project comprises ca. 60-70 thousand (as published in the Official Journal of the European Communities). The lack of precise indications as to the contents of this project (which documents should be translated) made it impossible to assess the exact timing and need for financial and human resources; moreover the situation has been changing ever since with the inflow of new information and the adoption of new legislation. By the same, the project team (as we will call the translation coordination unit) had to display a large degree of flexibility. However the main framework of the project remained unaffected.

2. The scope of the project

As mentioned before, the sources of the Office for Official Publications of the European Communities stated that the scope of binding acquis is assessed for 60 to 70 thousand OJ pages (what gives approximately 150 thousand standard pages). This large amount of pages should be translated into Polish within several years, as previously planned till mid 2003 (6 months before the date of accession to the EU). The priority has been set as to translate primary legislation, the list of which has been transmitted by the European Commission to Candidate Countries, followed by secondary legislation. At former enlargements, candidate countries (now Member States) received lists of acts to be translated. In the case of the present enlargement, no indication was furnished, the translation concerned the texts mentioned on screening lists A and B, Directives, Decisions and Regulations.

At present a common database for all candidate countries has been offered as a recipient for translated text. The possibility to check whether other countries decided to translates a given text might also serve as indication. However concrete information will be transmitted to the Office in Poland with the preparations for publication (indicating what documents will be published in a new special series of the Official Journal).

3. The structure of the translation unit

The structure adopted by national authorities competent for the translation of acquis communautaire into national languages is also related to budgetary constraints as all financial burden of the project is borne by the national budget. The solutions adopted by different countries can be generally classified into three groups:

1. *in-house translation unit* –translations are done by in-house translators, chacked by revisers and terminologists. This solution permits also to introduce computer-assisted translation and terminology management;
2. *revision unit* – translations made by free-lancers are revised by in-house revisers and terminologists. In this case the use of computer tools to translations implies either pre-translation activity lead by the unit or equipping free-lancers with the necessary software, what might prove very costly. Terminology consistency is checked at the stage of revision;
3. *coordination unit* – both the translation and the revision are carried out by free-lancers (usually public tenders to different operators), the unit performs only coordination and monitoring

functions. In certain cases it also takes part in settling major translation problems, most often related to terminology issues.

The choice of one of the above-mentioned solutions influences in great extent the adopted workflow, patterns of quality assurance and the necessity to develop terminological resources.

In Poland the structure of the revision unit was adopted. At present it comprises 20 persons and the structure is the following: 2 persons management, 1 secretary, 1 database manager, 2 terminologists and 11 linguistic revisers.

4. Workflow and organisation of the project (EU secondary legislation)

4.1. General workflow

Three main stages of the whole process of preparation of translation are the following:
1. translation;
2. revision;
3. archiving.

The translation stage is preceded by planning which is formalised and takes the form of yearly translation plans adopted by an organ of higher administration (Committee for European Integration). This plan is created in cooperation with line ministries and central administration bodies and its implementation is monitored by the Translation Department. The important information concerns the data of acts which translation is considered as a priority for the year in question, and financial resources – whether they are sufficient for the realisation of these plans. Gathering all priorities of different ministries permits to control the pace at which the progress in terms of translated pages is made and to undertake measures in order to accelerate the project if need arises.

Although the unit was set up in 1997 the lack of financial resources for the translation of aquis communautaire in general (staffing, equipment, etc.) did not permit to chose freely the optimal solution from the point of view of quality. Translation is a complex process and the lack of time is one of the main factors which has a very negative impact on the quality of the product (translation itself). The fact how important the issue of translation is, became more evident to the decision making authorities only in 2000 and only then the Department in its present form was created. The lack of time and money did not permit to adopt an in-house translation and revision system and the method of revision unit was continued.

The general stages of the project illustrates the following figure:

planning   **yearly translation plans**

I translation
- 1. **original in English**
  (for graphical elements paper versions; also in French or German if need arises)
⇩
- 2. **translation into Polish**

⇩

II revision
- 3. **revision / terminology**
  - consultations with experts
  - publication of collections of glossaries
  - up-dating terminology database
  - terminology help-desk

⇩

III archiving
- 4. **archiving in database**
  - Internet
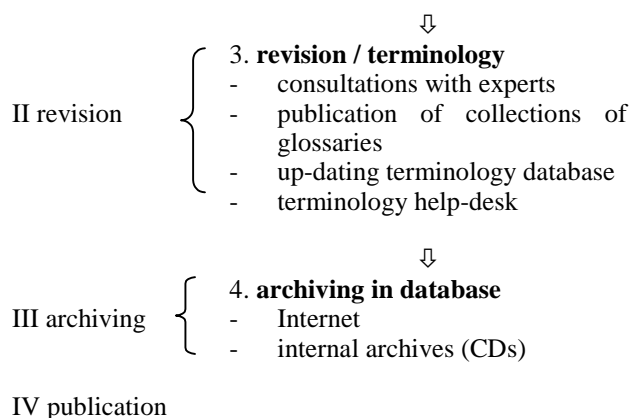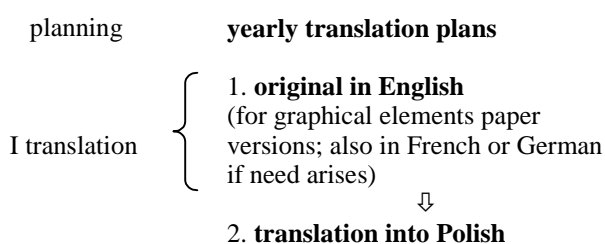 - internal archives (CDs)

IV publication

Figure 1: General workflow.

As the figure above demonstrates it, lists of texts for translations are prepared on the basis of translation plans and translation are contracted usually to translations agencies. Texts in Polish are submitted to the Department in batches and part of them is revised within the time span laid down in the contract. Corrected texts with a file of most common errors are returned to the translation agency which is supposed to correct all of its texts taking into account these errors. It is also necessary to mention that this "education" process of the contractor is very time consuming but the time left does not permit to check all texts. After the end of a contract texts are stored in a database and become accessible to the public. Revised texts (finalised) are stored in a European Commission database acting as recipient for candidate countries translations of European law. All texts are also regularly backuped on CD-Roms.

4.2. Planning; translation

As the project is entirely financed out of budgetary means, there is a necessity to carry out calls for tenders for translation services. The short time and the amount of pages do not permit individuals to provide these services. The conditions of the tender require not only long and satisfactory experience in translation of Community texts but also long-term cooperation with experts from different fields who would check the proper use of correct terminology. The originals are submitted mostly in English. The preparation of these texts is very time consuming as the EU electronic database of legal texts (Celex) does not contain graphics, tables and other elements, which demand some special formatting. These have to be handed out to the contractor in paper, usually copied from microfiches or lately from a set of CD-Roms, which was transmitted to the project team in December 2001.

4.3. Revision

The stage of revision is carried out by in-house linguistic revisers. All revisers have graduated from linguistic studies, have experience in translation of legal texts and large majority has accomplished postgraduate studies in European integration. Unfortunately not all texts received from the contractor (translator) can be immediately revised by the Translation Department (the same factors – time and the amount of texts). Usually

ca.10% to 20% of each batch is immediately revised and an error message is sent to the translator. This control revision is carried out on subsequent batches of texts until the end of the tender. The question might arise what happens with the remaining texts if only 20% of the whole lot are corrected. The time we have does not permit to revise all texts. Therefore, the translator is due to treat the remarks sent to him (error message) as a general indication and screen the remaining texts from the point of view of these mistakes.

The other two questions to be answered are terminology consistency, consistency among revisers and quality assurance within the Translation Department. The system adopted in order to meet high requirements of quality of translated texts requires also internal control of revision remarks and coherence of measures undertaken by revisers. Before storing the texts in the final database some revised texts are submitted to supervision. The remarks of the revisers gathered by the head of the project team are next transmitted to all its members. This permits not only to maintain a certain level of work but also to unify the aspect how far-reaching the corrections should be. Meetings of the revision staff as well as common terminology database are the tools for assuring uniformity of the adopted approach, as well as communicating to other revisers certain decisions, usually of terminology character.

The revision which is carried out in the Translation Department is of linguistic character. Texts are being carefully checked through and compared with originals in English and if doubt arises, in other official languages – usually French or German. The main aspects which are checked are the following: consistent use of adopted terminology, uniformity of cited titles of acts, names of institutions, committees etc., formatting (according to indications transmitted by the Office fot Official Publications of the European Communities); exact use of phraseology, especially concerning highly repetitive last parts of secondary legislation acts.

One might clearly say that the unit plays mainly the role of a quality control organ in respect of translation agencies. It also consults doubts which arise during revision and transmits recommendations in respect of terminology to translators.

4.4. Terminology

Consistent use of appropriate terminology is one of the key issues for quality translation. The said consistency is important only within one text but also among many texts which all together should form a coherent corpus. Efforts made in order to maintain terminology consistency concern mainly the following fields: legal terms, subject-related specialised vocabulary, phraseology, names and abbreviations.

The basic rule which should be followed is using the same terminology for the same concepts, in case of doubt always checking other language versions (adopted solutions). The language of EU legal acts is very specific – in general terms rather coherent, sober and mercilessly "productive": giving birth to new terms in order to name a notion which is being introduced. As a country with a different legal system, Poland faces the challenge to invent and generalise new legal terminology for notions

not yet in place. This is a very difficult task and should be solved on a case by case basis.

Three different situations can be encountered: when understandable equivalents exist in both source and target languages and their legal meaning is the same (1); when at first sight equivalents exist but their legal meaning in the EU law and Polish law is not exactly the same (2), when the terms which seem to be the appropriate equivalents already have different legal meanings both in EU and in Poland (3) and finally when there is no equivalent for a notion applied in EU legislation (4).

1.

 EU term = Polish term (same legal meaning)

2.

 EU term ≈ Polish term (different EU and Polish legal meaning)

3.

 EU term ≠ Polish term (two different legal meanings)

4.

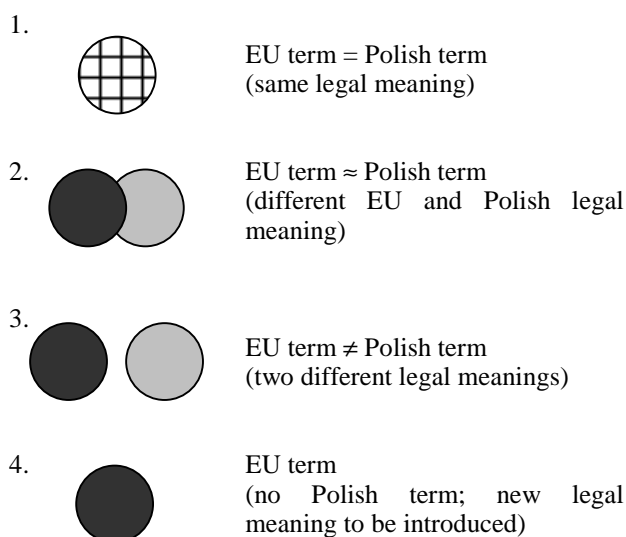 EU term (no Polish term; new legal meaning to be introduced)

Figure 2. Translation of terms between two different legal systems.

The decision indicating which is the situation we are dealing with or whether the lack of clarity in EU documents is accidental or made on purpose, is very difficult and responsible. As there is not legal revision scheduled in the workflow, the obligation to carry out consultations mainly with lawyers is the task of terminologists.

It is also important that official Polish terminology approved by the Translation Department is generalised and eventually used as recommended by vast public and all those who deal with European integration matters. To that end a general terminology database is accessible on the Internet[1] and glossaries of terms are published and distributed among translators.

4.4.1. Workflow; terminology help-desk

Problems which were encountered mainly at the stage of revision required an uniformed solution-finding approach. The workflow adopted currently is the fruit of experience rather than a decision taken ex-ante. The number of terminologists varied and so did the priorities and the results of terminology work for example in terms

---

[1] www.ukie.gov.pl is the web address of the Office of the Committee for European Integration (only in Polish). The link to the term-base and other resources can be found under http://www.ukie.gov.pl/uk.nsf/MD/120?edit&t=Tlumaczenia&c=120.

of the numbers of entries. At present the Translation Department has two terminologists.

In order to see clearly through the multiplicity of tasks which together can be talked about as terminological activity, we discern for this purpose three types of activity: 1. terminology help-desk (current problem solving and consultations), 2. uploading and updating terminology database, 3. compiling and publishing of glossaries.

The function of terminology help-desk is by far the largest and in daily activity it is the most important. Terminological problems which can not be solved by traditional ways (dictionaries, texts in target language, other databases) by revisers are transmitted to terminologists in order to find the right solution. Terms are searched for mainly in texts originally written in Polish on the subject. Proposed equivalents are often consulted with experts (usually employees of the ministry competent for the domain in question) or lawyers, if the doubt has a clearly and purely legal nature.

It is important to see the absolutely vital and central role which play terminology resources in this project. Network access to common term base is the only way to assure that the same terminology is used by all revisers and that they profit from findings of other revisers/ terminologists. Unfortunately the use of translation memory tools proved to slow and inadapted for revision purposes. This would most certainly assure consistency in a higher extent especially in respect of phraseology and repetitive parts.

4.4.2. Terminology database

Terms checked and approved within the Translation Department are stored in a terminology database. Several databases are used, for example, a global database with unchecked terms in majority (ca. 16 500 entries), export database comprises ca. 1200 terms of general legal nature. For internal use of revisers there are also thematic glossaries (for example on the common organisation of market in milk).

Terms are entered by terminologists into the database upon the proposal of revisers. The tool used for terminology management is Trados MultiTerm. All Candidate Countries have been sent by the Commission an empty shell – a template in order to preserve the model of the entry practised in the EC Translation Service and avoid any compatibility problems in the future. However the majority of the countries had to implement some changes in order to adapt this model to particular need of the unit. The model of an entry in our terminology database is the following:

---

system fields:
- user name
- creation date, change date
- entry number

user fields:

fields applying to the whole entry:
- documentary source
- project (pick list); name of a given collection of terms
- subject (pick list)

---

term (French, English, German and Polish)
fields applying to the terms:
- reference of the term, definition + reference, note + reference

---

Figure 3. Set of fields in the terminology database

Not all fields are filled in, usually only the entry terms and the reference. In addition it was necessary and very helpful to enter information confirming the reference of the Polish term. The fact that the Polish term comes from the same document as the terms in other languages is not sufficient to be assured about the quality of the term, e.g. whether it was checked, where it was found or who confirmed its actual use. There is also a habit of mentioning the date when the term was consulted as it may not be the same as the creation or change date of the entry. The field relating to a collection, glossary etc. permits fast and easy filtering and offers the possibility to export a subject-related set of terms.

During the revision work many temporary term bases are created which are next imported into the central database. All revisers have read-only rights in order to avoid mutual overwriting of terms and entries. All operations are handled by terminologists.

4.4.3. Glossaries

One of the drawbacks of the term-base which is accessible on the internet is that not all translators have free or easy access to these resources. Moreover, the distribution and use of lexicographic resources in electronic format is not always perceived as user friendly in Poland. Publication (in paper) of terminology collections permits to reach even those for whom the digital barrier is too significant.

Most of the glossaries which were published by the Office are European Commission's glossaries on different subjects. In early 1998 the EC Translation Service sent to the competent translation coordination units in the candidate countries a set of several collections of terms, recommending them as a good basis to start and continue with. Up till now 4 glossaries have been published, focused on the internal market, regional policy and economics and monetary issues.

Apart from glossaries, the Translation Department elaborates also other resources which it finds useful for translators. At the beginning of the project a short manual was preparedwith indications concerning the translation of EU legal acts into Polish. The referred to manual, called simply "Guidelines for translators" contains information on certain decisions related to Polish form of EU secondary legislation. In subsequent editions certain things have been altered, other added for the sake of clarity and due to the specificity and requirements of the Polish language. A large part is occupied by forms of acts in English, French and Polish, followed by examples of translated acts. These forms of acts contain also typical phraseology, which especially at the beginning and at the end of the document are highly repetitive and are almost always there unchanged. It is important to check the adequate use of these phrases, as ideally one and the same phrase in the source language should have only one and the same equivalent.

4.5. Archiving

Archiving is a stage common for all translated documents (primary and secondary legislation, case-law, etc.). Information on translated texts, translation plans and translations in course is stored in a textual database, for internal use of the Department. Files are stored on a server, accessible to all revisers. The data stored with the translation (such as date of the entry, author of the translation – competent line-ministry or usually the agency), English and Polish title, Celex number, number of Official Journal pages, etc. permits to provide reliable statistical information and to assess progress.

Another tool made available to this end this time by European Commission, is a database called CCVista[2] - named after SdTVista, an internal tool of the EC Translation Service storing all translations which have ever been done by this Service (Jank 2001). The new feature which has recently been added is a dynamic link with another database (Progress Editor), which primarily was used to monitor the progress of the harmonisation of legislation of candidate countries with EU legislation. This last database has however very well developed reporting facilities and it will now be possible to report on the contents concerning translations (input of data via CCVista) for example by thematic chapters, by chapters of Directory of Community Legislation in Force. CCVista itself is a simple repository for translations in all languages of candidate countries and does not offer all precious features which are offered by its original version in the EC Translation Service. Texts uploaded into that database are finalised translations.

5. Translation of EU primary legislation

The most appropriate way of starting this project would be to begin all translations by primary legislation. However as the follow up of the project was unknown and due to the fact that when state based translation coordination unit approached the problem, already existed multiple different translations of primary law, considered as the most important part of acquis, often used in raw translations for academic purposes. The issue of copyright arose as all translations which will be transmitted to the EU organs upon enlargement should be the property of the Office without any claims of third persons against them.

In the Polish case, and this situation is similar in other countries, secondary legislation was contracted for translation before having finalised the primary legislation. Although there were several versions of the Treaties (they are the most important pieces), they only served as indication. They could not be exact signposts for the vocabulary and provisions which are recalled in the secondary legislation. Primary legislation is currently being finalised. The work upon the new accession treaty has already started within the Council. As the contents of the accession treaties does not vary considerably and for the sake of consistency translation memory resources are

---

[2] The database is accessible via Internet: ccvista.cadmos.carlbro.be. Files are downloadable in Word format (free of charge).

currently being built up. The finalised treaties are post-aligned and stored in the translation memory. Missing documents and acts in course of revision are checked for consistency by multiple functions of the Translator's Workbench. The drawback observed is a general problem with fonts translation and lack of diacritical signs proper to languages on the candidate countries. A clean-up macro is not the best solution, as it leaves doubts as to the proper spelling of segments contained in the translation memory.

The timing for primary legislation is very precise. All documents mentioned on a list drawn by the EU legal services should be finalised by mid 2003.

6. Translation of ECJ case-law

The last stage of the project consisting in translation as well, is the translation of case-law of the Court of Justice of European Communities (referred to as ECJ). All candidate countries obtained a list from the Court with a choice of ca. 800 judgements. The total volume of texts amounts up to ca. 15 thousand pages.

The difference of approach adopted in this case in that there are practically no translations of case-law yet. The translation of case-law can be treated as a separate project and that is why it will only be briefly mentioned in this paper. In general the text which should be translated for each judgement is the resume, key words and the judgement itself. According to the recommendations of the Court the texts should be translated from French (French is the official language of the Court) and they should follow chronological order. The fact that there are highly repetitive parts can mean that translation memory facility and pre-translation might prove useful.

Currently there are significant problems with financing of this part of the project. The Translation Department is preparing a Polish version of a manual on the judgements of ECJ which will serve as reference point and will be binding for translators. A database of key words is being built (key words are taken from all listed judgements). This database in French, English and Polish is not only supposed to help translators and avoid multiple translations of the same notions but also permit to import these notions to a database of the Court (one of search criteria).

7. Publication; end of the project

The publication of EU legislation in languages of the new member states, to join in by 2004, is the obligation of the Office for Official Publications of the European Communities. The Office wishes to start preparing itself and the texts before the date of accession, and the texts will be transmitted by candidate countries as from 2003. The issue of publication is very complicated as such volume of texts has never been issued simultaneously in 10 languages or so. The additional difficulty is that texts will be published by thematic chapters and the lists of documents to be published might be transmitted to the candidate countries very late. Taking into account the budgetary planning procedure it will be hardly possible to foresee the amount of money necessary

for any additional translations lacking or fixing other problems. This stage for the moment remains a mystery and it will demand a lot of flexibility and fast reaction in order to keep up with the planned pace of publication.

The above-mentioned publication is addressed to new citizens of the EU. Until that time there is a general need for public administration and business sectors to access translation in their own language. To this end serves Internet accessible full text database (aforementioned web pages of the Office). The texts accessible in that database are only revised translations, non-revised translations can be obtained upon request addressed to the Translation Department.

8. Closing remarks

Various approaches adopted by the candidate countries for the same goal – translation of acquis comunnautaire into their national languages – displayed the complexity and the constraints which are usually encountered in large public administrations. The team of the Translation Department also gained knowledge in course of the project as it should not be forgotten that EU legislation with its specific vocabulary was until recently an unknown domain in Poland. Certain decisions which were taken at the beginning had to be changed, other ones were good and served as basis for subsequent decisions. It is true particularly for terminology where a completely new terminology had to be build, invented at times.

As the projects develop and candidate countries are closer and closer to the day of accession, several questions and comments arise. One of them is the fact that there was too little contact among translation coordination units of different candidate countries and that this has led to multiple efforts, repeated in each country (like scanning certain documents which are to be translated by everybody, building up databases of key words, where the common language structure on French and English could be shared etc). The other question which comes up out of practical reasons, is why a unique model of translation coordination unit was not proposed by the European Commission with a tight schedule for progress in that matter. This would most certainly permit to avoid doing the same job many times in different countries and it would assure, among other, a common terminology database ready for input to Eurodicautom as well as a concrete set of act which would eventually be translated and published. On the one hand a unified structure would render the whole coordination much easier. On the other hand, the application of various methods to one and the goal permits to assess which was the best solution.

References

Jank, D. (2001) Structure and operation of the EC Translation Service, *Procedings of the conference on the translation of acquis communautaire*, Warsaw 2001.

McCluskey, B., (2001). Respecting multilingualism in the Enlargement of the European Union – the organisational challenge, *Terminologie et Traduction*, 2/2001,   Brussels.