

Generic Text Interface (Canevas/GTi): Producing and publishing structured legal documents at the Court of Justice of the European Communities

Brian Porro, Jean Acheroy and Florent Thouvenin

Translation Directorate

Court of Justice of the European Communities

L-2925 Luxembourg

00 352 43033435

brian.porro@curia.eu.int

Synopsis

GTi allows the author or an analyst to access all background documentation relating to direct or indirect references inserted in his document. The system aims to track - at the authoring stage - and subsequently retrieve all the background information needed by lawyer-linguists by producing an annotated version of the document containing directly clickable references to the background documentation. Using this multilingual reference system, a translation matrix is generated based on the author's document. Each matrix contains a specific linguistic version of all inserted citations as well as bilingual links to the original documents, thus allowing direct verification of the translation context for each citation.

I. - Background

The rules on the use of languages

1. The rules governing the use of languages at the Court of Justice and at the Court of First Instance of the European Communities are a special feature which distinguishes them not only from other national courts but also from other institutions of the European Union.
2. Proceedings before the Court of Justice and the Court of First Instance may be conducted in any of the languages defined as official in the legislation governing the work of the Community Courts. This means that, from the moment when the application initiating

proceedings is lodged at the Registry, the parties and the Community Courts will use throughout the proceedings (at least for the written submissions) the language indicated as the language of the case in accordance with the Rules of Procedure.

3. Moreover, the uniform application of Community law requires the dissemination of the case-law in all the official languages of the Union, since all citizens of the Union are entitled to have access to it as a source of Community law, in the official language of their country. This dissemination is achieved not only through the publication of the definitive (and only authentic) version in the Court Reports, but also with the instantaneous online reports of the Curia website.

The practical difficulties of the language rules

4. The language rules were established in order to provide untrammelled access to Community justice to those persons subject to the jurisdiction of the Court. In principle, the person concerned uses his or her own language, the Community Court gives a ruling in that same language and publishes it in languages understood throughout the European Union. However, these rules involve real difficulties in practice. Obviously, not all the members of the various chambers of the Court who come from various Member States are expected to master the language, either passively or actively, of each case in which they are involved.

5. Since its creation in 1952, the Court of Justice has imposed a strict rule of confidentiality of the deliberation. The decision adopted at the end of the deliberation is signed by all the members of the chamber taking collegiate responsibility, irrespective of the complexity of the debates or of whether the final version of the draft decision was adopted by a majority or unanimously. Those strictly technical reasons led the institution to decide on a single language, which the members of the chamber are expected to know thoroughly, in which the draft judgment is drawn up by the Judge-Rapporteur, the deliberation is conducted and the draft judgment adopted.

Organisation

6. The way in which the institution has organised its work has stood the test of time during more than half a century of activity despite the steady increase in the number of official languages (4 on foundation, 20 after the next enlargement of 1 May 2004).

7. The pleadings submitted by the parties are translated into the language of deliberation and into the language of the case (where they are different).

8. The opinions of the Advocate General, the judgments and orders are translated into the language of the case for the purpose of notifying the parties and into all the official languages for the dissemination of the case-law.

II - The role of the Translation Directorate in the application of the language rules

9. That organisational framework represents three very important challenges for the Translation Directorate, the institution's legal translation department:

- translation must be turned around as quickly as possible to avoid the absence of translation within a reasonable time-limit being the cause of delay to the proceedings;
- the quality of the translations has to be impeccable, since any error could impair the sound administration of justice of the Community and the uniform application of Community law;
- when the language of the case is different from the language of deliberation (which is the most frequent situation), the judgment itself (the one signed by the judges and notified to the parties) is not the text discussed and adopted by the judges in the language of deliberation but the text provided by the Translation Directorate: by a legal fiction, it is when the judges sign the “translation” that the text in the source language becomes a translation and the version in the target language becomes the original.

10. The requirement that all language versions be of the best quality, in addition to the imperative need for faithfulness of the translation, has always required the lawyer linguists translating the case-law to make the most enormous efforts in terms of research: a “mistake” can consist not only in mistranslation or erroneous adaptation of terms into another language but also in making the wrong choice of the source of a legal citation (explicit or implicit). Indeed, any divergence in the choice of the source may introduce wording into the target language which deviates from the original intention of the originator of a text. The difficulties involved in complying with those requirements are not made any less by the strict confidentiality surrounding the deliberation.

11. The following example (paragraph 13 of the judgment of the Court of Justice of 16 January 2003 in Case C-398/99 Yorkshire Co-operatives Ltd) serves to illustrate the point:

“It must first be borne in mind that, according to settled case-law, in the context of the cooperation between the Court of Justice and the national Courts provided for by Article 234 EC, it is solely for the national Court before which the dispute has been brought, and which must assume responsibility for the subsequent judicial decision, to determine in the light of the particular circumstances of the case both the need for a preliminary ruling in order to enable it to deliver judgment and the relevance of the questions which it submits to the Court. Consequently, where the questions submitted by the national Court concern the interpretation of Community law, the Court of Justice is, in principle, bound to give a ruling (see, *inter alia*, judgment of the 15 December 1995, Case C-415/93 *Bosman* [1995] ECR I-4921, paragraph 59, and judgment of the 17 May 2001 Case C-340/99 *TNT Traco* [2001] ECR I-4109, paragraph 30)”.

12. By referring to the *Bosman* case and the *TNT Traco* case in brackets, the Court gives the reader the main legal precedents underpinning its reasoning. Each lawyer linguist translating the judgment into each official language should check whether or not this paragraph was taken verbatim from paragraph 59 of the *Bosman* case or from paragraph 30 of the *TNT Traco* case. If so, he should still determine which of the two sources was actually used in order to insert the appropriate citation, either the *Bosman* case or the *TNT Traco* case, within his or her translation which must, in addition be the text as it appears in his or her language version of the Court Reports.

13. This relatively simple search can take from 5 to 10 minutes for each target language with the standard means used by lawyer linguists. However, in view of the forthcoming enlargement of the European Union on 1 May 2004, the sum of the time devoted by nineteen lawyer linguists to research represents a minimum of two hours' work. The Translation Directorate could save this time if the precise source of the citation is furnished to the lawyer linguists either with the text to translate or with a draft outline of a translation. The scale of this saving increases in direct proportion to the level of difficulty of the research and the number of languages that must be dealt with. That economy of scale should not detract from the qualitative gain represented by such "horizontal" or "synchronic" consistency among the various language versions, which is essential to the uniform application of Community law.

III - Towards a multilingual production of the documents at the Court of Justice and the Court of First Instance

Background: examination of the "repetitiveness" element and tools for computer-aided translation (CAT)

14. During the latter part of the 80's, under the pressure of an increasing workload which was ever more difficult to handle, the Translation Directorate began to look at computing tools which might facilitate the task of the language lawyers and thus increase production as a whole without sacrificing the high level quality and consistency of translations. At that time, the term "repetition" was very much on everyone's minds in CAT circles: software based on translation memory began to have enormous success on the market. The fact that the software was based on an obvious idea did not detract from the revolutionary benefits it brought to language engineering: if a translated text contains sentences or other "translation units" already contained in previous texts which have previously been translated, the task of the translator would be made easier if a computing system (be it local or networked) displayed suggested wording from former translations for each new sentence or "translation unit" required.

15. Despite its attractiveness, the application of this idea using the software of the time entailed certain limitations which were not readily acceptable in terms of legal translation:

- although “translation memory” could display suggested translations based on previously-translated sentences to lawyer linguists, they could not provide direct access to the context from which the “translation unit” had been drawn. It is, as explained above, the context which determines the relevance of any item (from a whole sentence to a single word) to be inserted;

- because the content of the translation memories available in each pair of languages is not always the same, there is a risk that different choices might be proposed to lawyer linguists translating texts which are nonetheless identical. This possibility evidently undermines the “horizontal” consistency of the different language versions;

- translation memories were initially created with self-employed translators or translation agencies in mind, and were thus difficult for an institution such as the Court of Justice to manage, particularly because of the system of quality control which is employed. The documents are checked at various stages of production, involving the handling of successive versions, so that use of translation memories would have made such heavy demands in terms of resources needed to update versions and “clean-up” documents that the savings in translation would have been neutralised;

- finally, one last element militated against translation memories: because legal texts are characterised by longer sentences on average compared to those in non-legal texts, the “reusable” translation units are mostly textual segments which are longer than those which could be identified by the alignment systems which update translation memories. Thus, translation units which were in fact repeated were not necessarily picked up by the system. Studies carried out at the time using the corpus of documents to evaluate the “rate of repetition” of the translation units did not give decisive results for this reason.

16. In this context, any consideration of the concept of “repetition” had to follow another approach bearing in mind that, for 80% of the texts to be translated, the originator of a text and the lawyer linguists both belong to the same organisation. That meant that it was possible to handle “repetition” not at the stage of the translation of documents, but as early as the writing stage. That view was supported by the following observations:

i. repetition stems from the fact that the originator of a text uses existing texts in particular within the framework of citations required by legal reasoning and from the fact that they use

certain formulae because of the need for consistency, throughout the proceedings, with the legislation and the applicable case law.

ii. repetition is the result each time of the thought and research carried out by the originator of a text on behalf of the Institution. Consequently, this research belongs to the Institution itself and thus also to the translating service and to each of the lawyer linguists responsible for the production of the various language versions of each document.

Centralised system of references

17. That analysis means that establishing a system capable of being of help to translation should aim to achieve two objectives. First, the setting-up of a centralised system of analysis and references for documents sent by the originator of a text. Secondly, the creation of a work environment for the originator of a text to aid the drafting task and facilitate research, while leaving traceable marks of that research in the document. This system makes it possible to produce a draft (“the translation template”), using a special processing system, containing all the citations referred to by the originator of a text for the benefit of the various lawyer linguists.

18. Clearly, this approach can only increase the reliability of the “translation units” offered by the system, since they are drawn directly from the writing stage. Moreover, without knowing all the languages concerned, it allows the originator of a text to predetermine the wording of the citations and typical formulae not only in the language in which he is drafting but also in the other language versions.

Document memory and “logical” translation memory

19. In the course of considering such a system, it became obvious that the possibility of marking up text with standard structured languages available such as XML make it possible to envisage any document which used such a language, stored in a “document bank”, being automatically updated in real time: each time an XML document is saved in such a bank it would be possible to index it and ensure that the Court's document bank was continuously updated. This “document memory” would be of similar content in all the languages at all

times and, without the need for any further processing or handling, be immediately available not only to the lawyer linguists but also the originator of a text.

20. From that point of view, the natural segmentation of texts as a result of their being saved in XML format should make it possible to put in place a system of textual research which would be able to find, for a specific language, a certain sequence of words and identify parallel segments in other language versions without the need for a translation memory compiled on the basis of a “physical” alignment and without any of the shortcomings resulting from the large number of language permutations which require to be managed (380 from 1 May 2004). The expression “logical translation memory” thus refers to such an exploitation of a “document bank”. With all this in mind, the Document Bank project was set up in the Court aimed at attaining those objectives among others.

IV- Putting the chosen approach into operation: functional description

21. The Court thus decided to put underway the coordinated work necessary for establishing three discrete but consistent working environments for the drafting, translation and publication of all the documents produced within the institution.

The drafting environment

22. The drafting environment is the subject of the “Canevas 2” project, which is under the direction of a group of text originators set up for this purpose. Its aim is to enable the creation of a working context which, while being user-friendly, contains a number of restrictions necessary to ensuring continuity in production (drafting — translation — publication).

23. The goals pursued by this environment are the following:

- assisting creation of XML-format documents structured in accordance with the Vademecum of the Court of Justice and the Court of First Instance.

- entering content by means of a word-processing software (MS Word XP) retaining the original user functionality except in specific, well-defined situations where use of such functions could undermine the structural integrity of the documents
- drafting and reviewing tools designed to:
 - a) guide the way in which is entered parts of the documents whose structure is strictly regulated by the Vade-mecum;
 - b) supplement the standard Word revision tools to meet the particular requirements of the institution;
 - c) facilitate documentary research and drafting choices;
 - d) mark up the origin of text taken from elsewhere by means of research and drafting aid tools;
 - e) check and, if necessary, restore the structure of the documents.
- automating conversion of documents to XML format;
- making tools available for the storage and retrieval of documents in and from the document bank.

Creation of documents

- Structure of documents

24. From a strictly functional point of view, a document in Canvas format is composed of one or more “free” (or “unrestricted”) blocks in which the legal body of the text is written in

directly using word-processing software, and one or more “structured” (or “restricted”) blocks containing the procedural information of the case (subject-matter of the case, national origin, parties’ names, composition of the chamber, signature of the members, etc.). Procedural information is keyed into special Canevas forms in order to ensure that the structure of the document is correctly interpreted for subsequent conversion to XML format.

- Procedure for the creation of documents

25. When the application is started, a dialogue box (see figure 1) gives the user a range of options: in particular it can set up a full Canevas including the “free” blocks (albeit the conventional ones) of the document (the blocks containing the grounds, costs and operative part in the case of a judgment) and the structured blocks depending on type of document (Judgments, opinions, etc) and procedure involved (direct action, preliminary ruling, etc). In such cases, the Canevas is created taking into account the data available in the data base containing pending cases (Litige), maintained by the registries of both courts. The creation of a document may also be restricted to creating a Canevas without title page (introducing opinions or judgments). Finally, it is possible to create just the “free” part of the document which is suitable for subsequent inclusion in a Canevas.

CANEVAS 2.8.0-BUILD4 ✕

Jurisdiction Numéro (ex:168) Année (ex:02)

Numéro d'affaire: - /

Création

Création assistée du Canevas (Informations fournies par la base Litige)

Création manuelle du Canevas (Questions / réponses)

Création d'un Canevas sans page d'en-tête (Corpus seul)

Type de document

Informations supplémentaires (Création manuelle)

Type d'affaire:

Langue du document:

Avocat référent:

Chemin et nom du document:

Figure 1: Creation of "canevas"

- Handling the structured blocks (forms)

26. Even after they have been created, the structure, the format and the textual content of the structured blocks may be modified easily and at any time using a customised toolbar and specially-mapped keyboard shortcuts. Those tools make it possible to add or remove any of the discrete parts making up the document and to key in or change the text which should appear in the free-text area of each block (see figure 2).

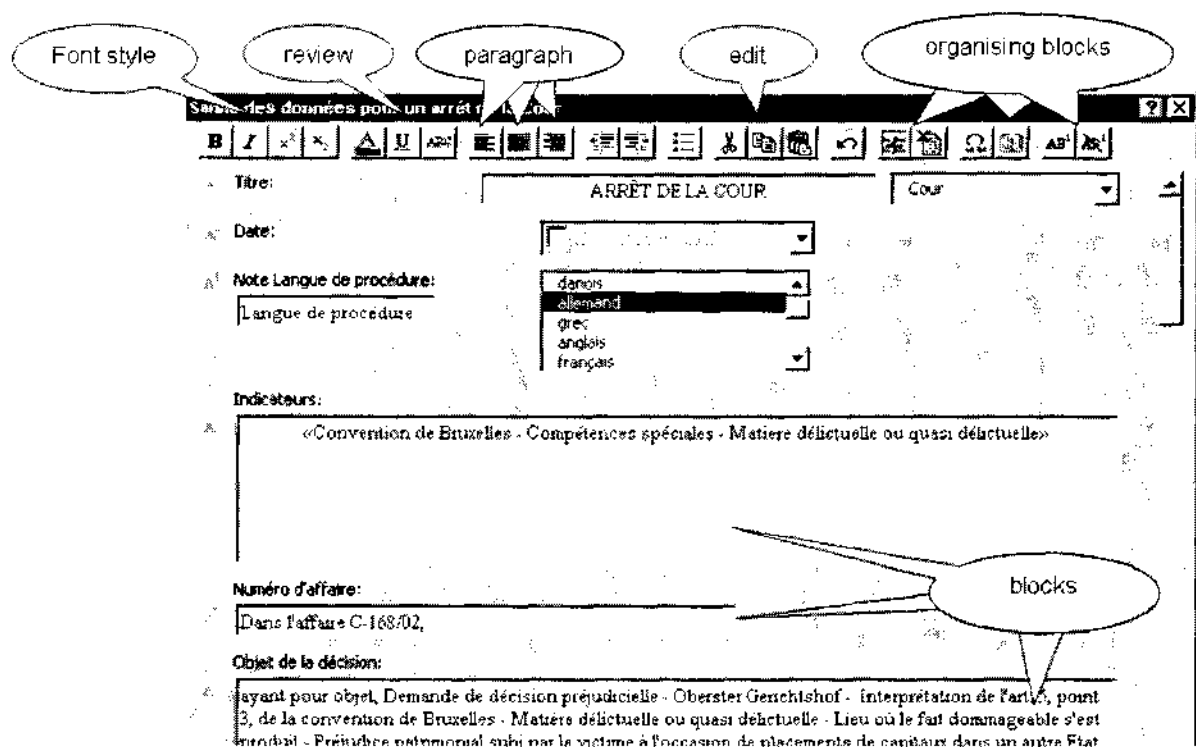


Figure 2: forms

- Typing into the “free” blocks (word-processing)

The user types the text up in the usual way, making full use of all the facilities offered by the word-processing software. The only constraint lies in the requirement to use the “styles” appropriate to the kind of document written. Those styles are managed using Word models or templates within the Canevas application. To make the choice of appropriate styles easier, they are labelled in the style pane with names which are as close as possible to the terminology normally used in the court. Furthermore, they are displayed according to frequency of use, so that the most-often-used styles appear at the top of the list (see figure 3).

As far as possible, the same style names have been used in the “templates” for the various kinds of documents in order to make copy and paste between two Canevas documents easier and safer.

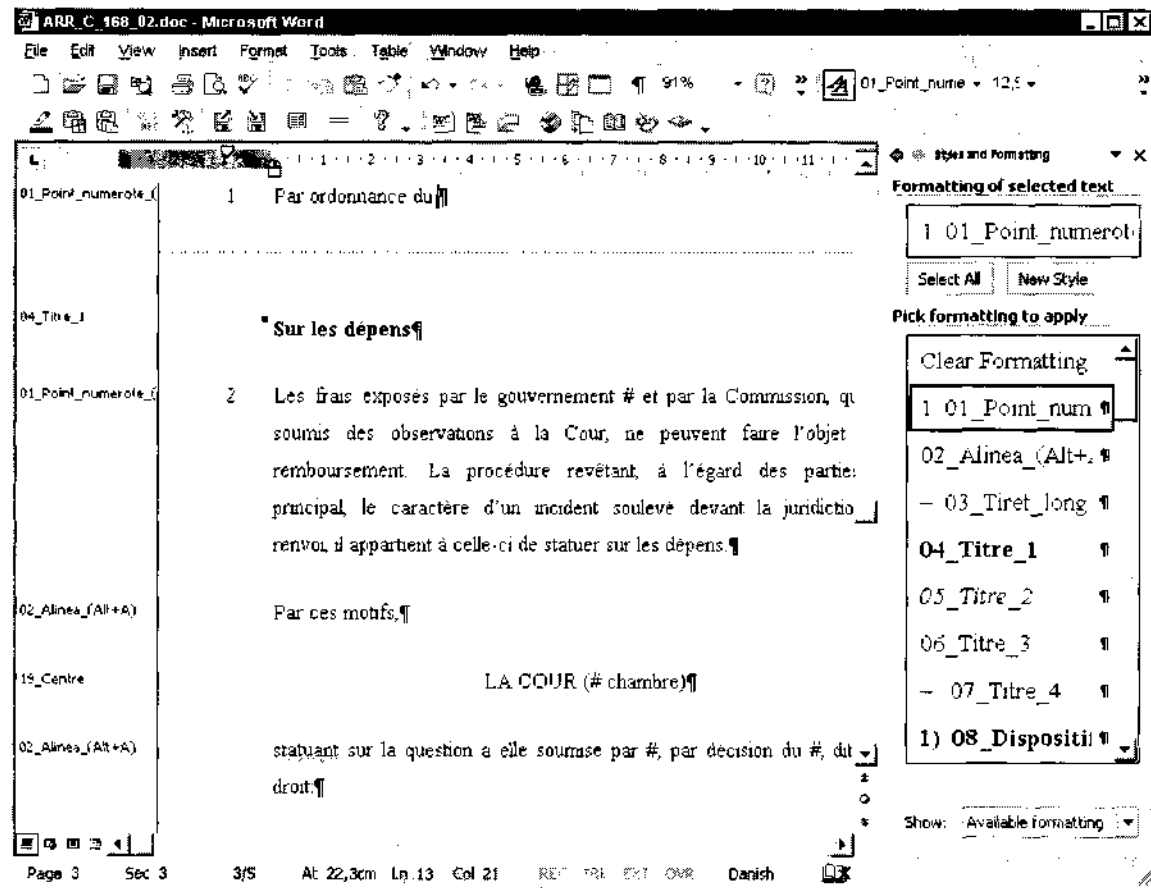





Figure 3: applying styles

- The Canevas tools

27. The tools presented by Canevas are grouped around four basic functions:

- 1) Access to Canevas from the word-processing platform


Functions available:


- a) create a new Canevas document 
- b) open an existing Canevas document 
- c) access to templates containing formatting styles 

- 2) Updating and checking (“verification”) the Canevas:


Functions available:


d) detect and repair the Canevas. 


e) insert inverted commas (according to level and language of the document) 


f) insert citations using the Canevas application ("copy using tools") 


3) Quality control of the texts

g) show modifications to document 


h) bold footnote numbers 


i) accept changes to footnotes 


j) delete character to left of the cursor 

k) change case of character to left of the cursor from lower to upper case 


l) insert ellipsis (...) 


m) activate "review" mode 

n) delete extra spaces 


o) insert apostrophe from character set 


4) Drafting and document search aids

p) consult the Celex legal data base 

q) consult the sentence library 

r) consult references to decisions of the Court of Justice and the Court of First

Instance published in the Court Reports 

s) search and insert citation using Canevas tools 

28. The last-mentioned group of tools play a central role in the multilingual system of production of the Court’s documents and thus merits more detailed description.

First, it must be pointed out that they have a double function:

- They give the originator of a text not only easy access to sources of information frequently used when drafting texts, but also the ability to insert into the document the result of their research keeping the structure of the target document while adhering to the typing rules under the Vade-mecum;
- They allow the originator of a text to leave a non-language specific trail of their research, making it possible at a stroke to insert automatically the relevant references in the translation templates subsequently generated in the different languages from the original document.

29. Returning to the example in paragraph 11 above, the system enables the originator of a text to send a document to the Translation Directorate which may look as follows:

«À titre liminaire, il y a lieu de rappeler que, conformément à une jurisprudence constante, dans le cadre de la coopération entre la Cour et les juridictions nationales instituée par l'article 234 CE, il appartient au seul juge national, qui est saisi du litige et qui doit assumer la responsabilité de la décision juridictionnelle à intervenir, d'apprécier, au regard des particularités de l'affaire, tant la nécessité d'une décision préjudicielle pour être en mesure de rendre son jugement que la pertinence des questions qu'il pose à la Cour. En conséquence, dès lors que les questions posées portent sur l'interprétation du droit communautaire, la Cour est, en principe, tenue de statuer (voir, notamment, arrêts du 15 décembre 1995, Bosman, C-415/93, Rec. p. I4921, point 59, et du 17 mai 2001, TNT Trago, C-340/99, Rec. p. I4109, point 30).»

Kommentar: (saisi ARRC-34099 point 30 CELEX)

30. The printed text (the one which will actually be published in the Court Reports), remains unchanged, but a comment will be added in the electronic “working document” version. This comment does not merely give accurate information as to the origin of the citation to other colleagues within the court but also contains a hyperlink which gives direct access to paragraph 30 of the judgment cited. At the same time, this hyperlink enables the analysis system and translation template generating system (described below at paragraph 35 et seq.) to identify the beginning and the end of the citation in the original document as drafted by the originator and to paste the passage thus defined in each of the requisite languages.

31. That result is obtained without the need for the originator of a text to take any further action. When the originator requires to refer to the case law as in the example below, he uses the “Citations help” tool which displays the following dialogue box:

Aide aux citations

Catégorie : Jurisprudence

Affaire : Cour - 340 / 99
Affaire ex:127 Année ex:99

Type de document : Arrêt

Exemples : 1 ou 1-3 ou 1,4,6-8

Liste des points : 31

Version linguistique : français

Insérer un marquage pour la traduction

Annuler Chercher et copier

Afin d'insérer dans un document Canevas le texte ayant fait l'objet de la recherche, il est nécessaire de placer le curseur à l'endroit désiré et ensuite, d'utiliser la fonction « insérer » de Word ou de cliquer sur le bouton « Coller depuis les outils » présent dans la barre d'outils du document Canevas.

Figure 4: citations help tool

32. Ticking the “mark up for translation” box has the tool not only search (internal servers and if necessary in Celex) for the text and the paragraph required and insert them at the correct place but also add a comment box detailing the source. Moreover, simply clicking on the link in the comment box allows the originator of a text to review at any time the case law or legislation cited in its original context.

- Sending the document to the Translation Directorate

33. When a document is ready for transmission to the Translation Directorate, the originator of the text changes its status to “ready for translation” by checking it into the document bank. The document is then converted to XML format.

- The translation environment

Pre-processing, analysis and generation of translation templates

34. It is to be expected that the level of use of automatic reference tools by the various originators of text will vary, particularly during the transitional period. Nevertheless, the Translation Directorate cannot make the management of translations conditional on an aspect of the organisation over which it has no control. It has therefore kept in place the system created in the second half of the 1990's in the Directorate which is based on a centralised system of analysis and referencing of documents. The analysis and referencing unit has the task of carrying out once-for-all research for all the required language versions where the originator of the text has not marked up his sources. If necessary, it can also remove or correct erroneous electronic references inadvertently left by the originator of a text in the course of successive modifications of the document as it goes through deliberation.

Allocation of the work of analysis

35. A request for translation is sent to the Central Planning unit of the Translation Directorate. Central Planning assigns the work of analysis to a lawyer linguist in the analysis and referencing unit. The allocation of this piece of work, via the translations management database (Suivi), sets off an automatic procedure which allows the system to link the metadata of the translation request to the metadata of the document which is stored in the document bank. Once matched up, the system sends the analyst an e-mail containing a hyperlink to the XML version of the document to be analysed.

The analysis desktop

36. The interface displayed by the “analysis desktop” is a clear workface which can be customised (see figure 5: the analysis and translation template generation system). On the desktop the user may display documents, a menu bar, a tool bar, and a control panel to show all the references making up the background file.



Figure 5: the analysis and translation template generation system

Pre-processing

37. Accessing the XML document stored in the document bank in order to analyse it activates the pre-processing of the document and the creation of a background file. The result of that process, which takes only few seconds, may be seen in figure 6.

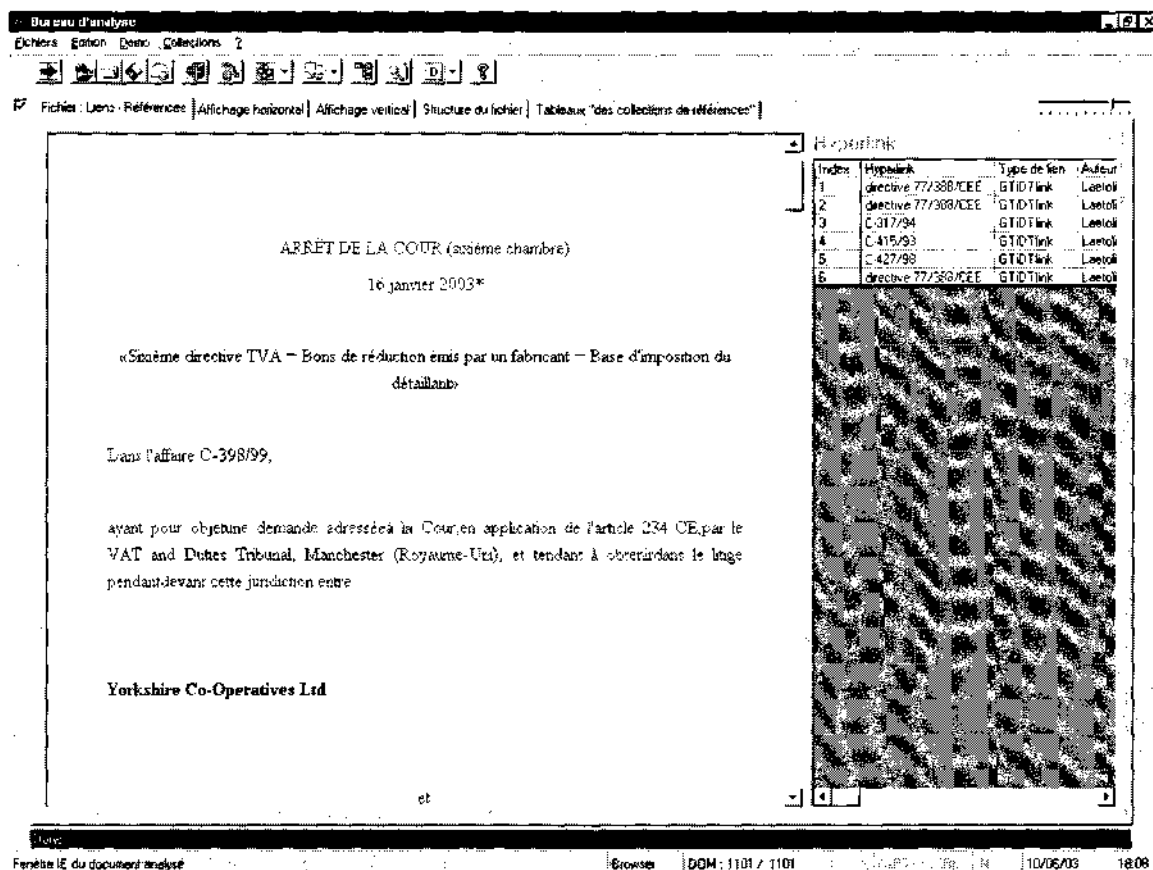


Figure 6: Pre-processing; standard phrases

38. The pre-processing program identifies the standard phrases contained in the introductory page (underlined) and differentiates them from procedural information. The analyst can then estimate the level of success of the program for translation of that page into every target language. If necessary, he can ask the phrase-handling unit to add, correct or remove phrases or sentences.

39. The pre-processing program also identifies expressions which could be references to Community legislation or case law, such as “Directive 77/388” and “C-317/94”. It changes those references to hyperlinks so that the analyst may go directly to consult that legislation or case law.

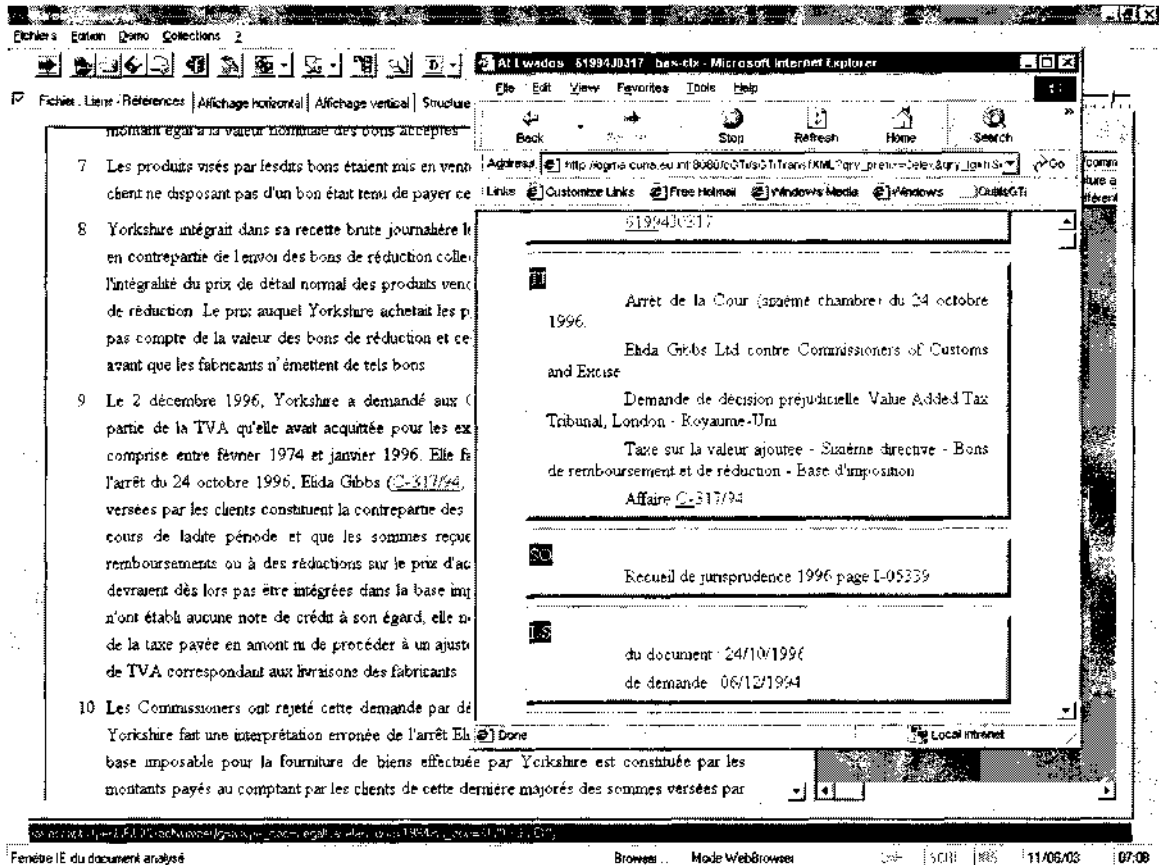


Figure 7: hyperlinks added

40. Finally, the pre-processing program identifies the citations made by the originator of the text using Canevas tools and highlights them for the analyst to verify and, if necessary, supplement them with different references for languages in which the documents referred by the originator do not exist in the document bank or in Celex (this is especially the case when the citation comes from a provision or case-law prior to the accession of a Member State).

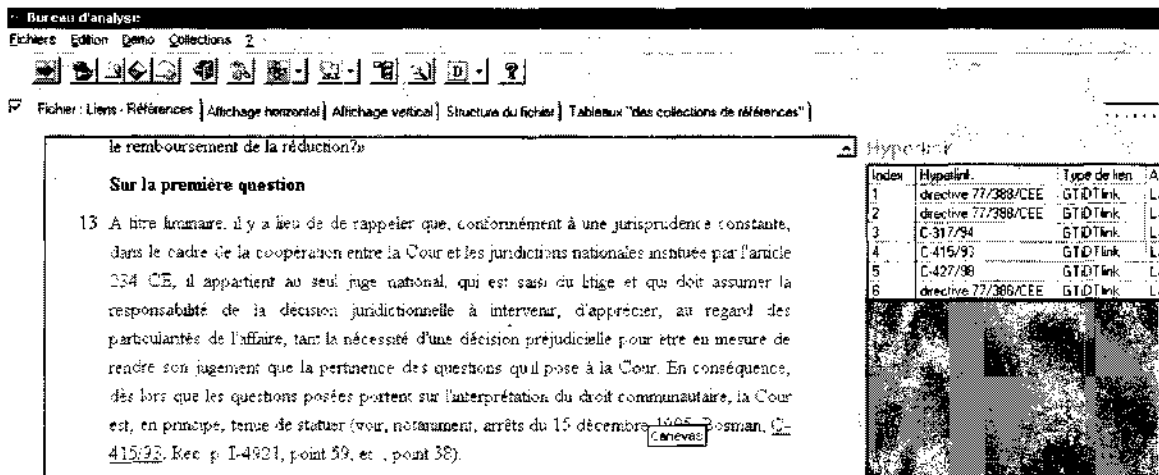


Figure 8: Pre-processing - trace of citation inserted by originator

41. When pre-processing is completed, the analyst can call up the “background file” (in this case the series of references) in a pane. Clicking on one reference in the pane gives immediate access to the precise spot in the text where it is located.

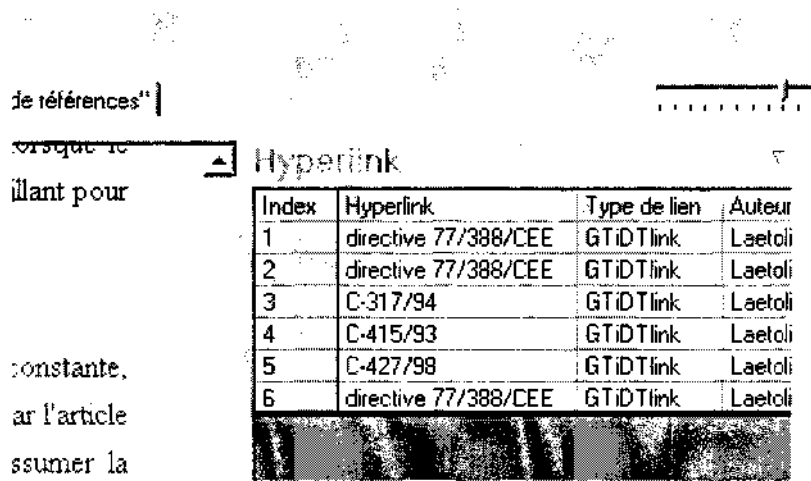


Figure 9: Pre-processing - set of hyperlinks

The analysis

1) Enrichment of the document

42. Next, the analyst can move throughout the document in order to insert manually, using the tool bar, any comments or capture commands.

43. Returning to the earlier example, if he should notice that paragraph 14 contains a citation which has not be marked up by the originator of the text to enable a multilingual capture, he can activate the capture tool and insert a citation command using a function similar to the one available from the Canevas tools.

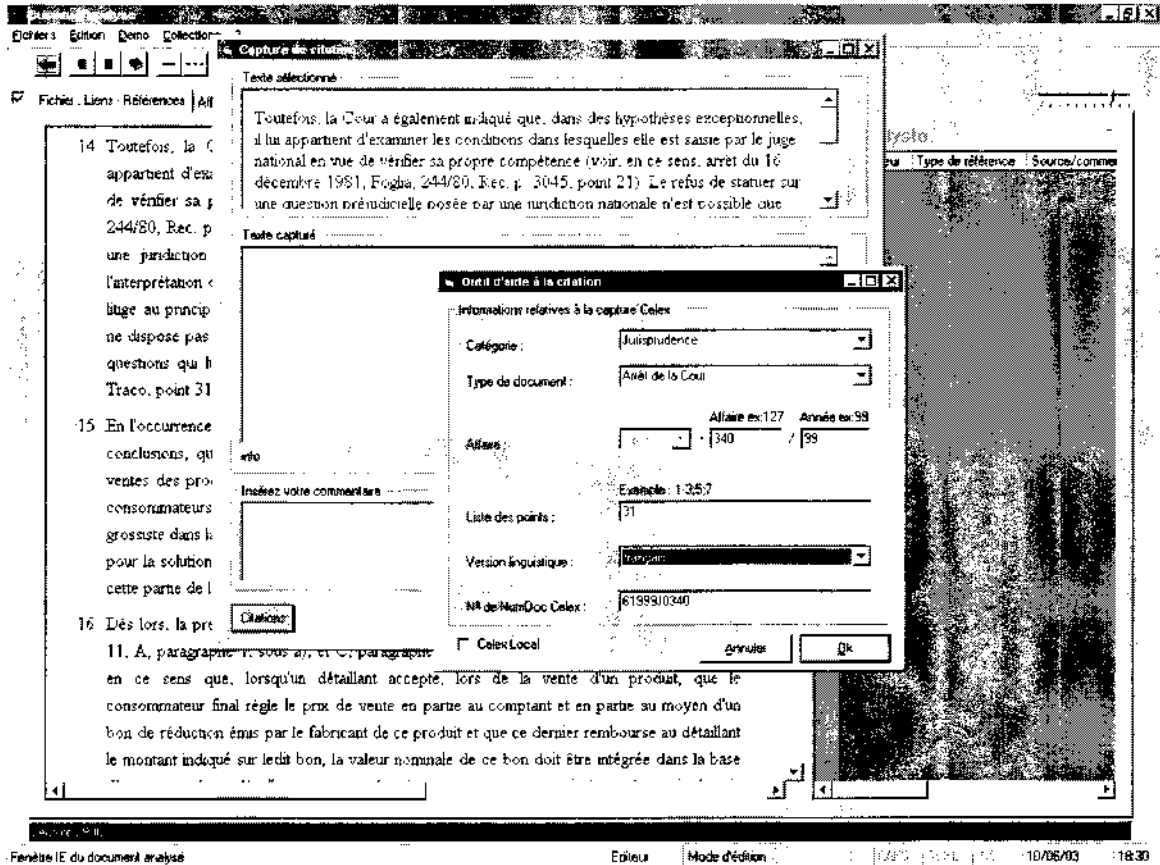


Figure 10: Analysis - insertion of a capture command

After having carried out the necessary research, the capture tool displays the source text of the citation to enable the analyst to compare it with the text shown in the document to be analysed and to add comments which he considers relevant for the guidance of the lawyer linguists.

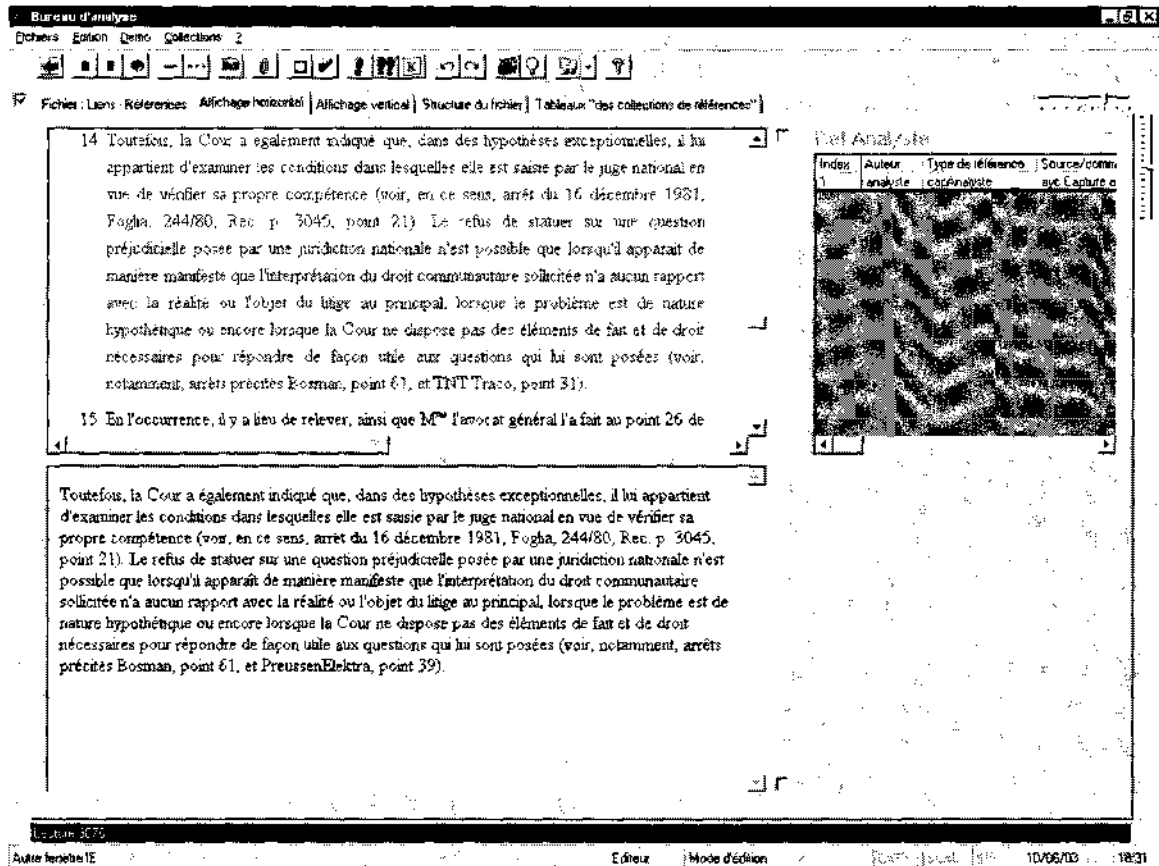
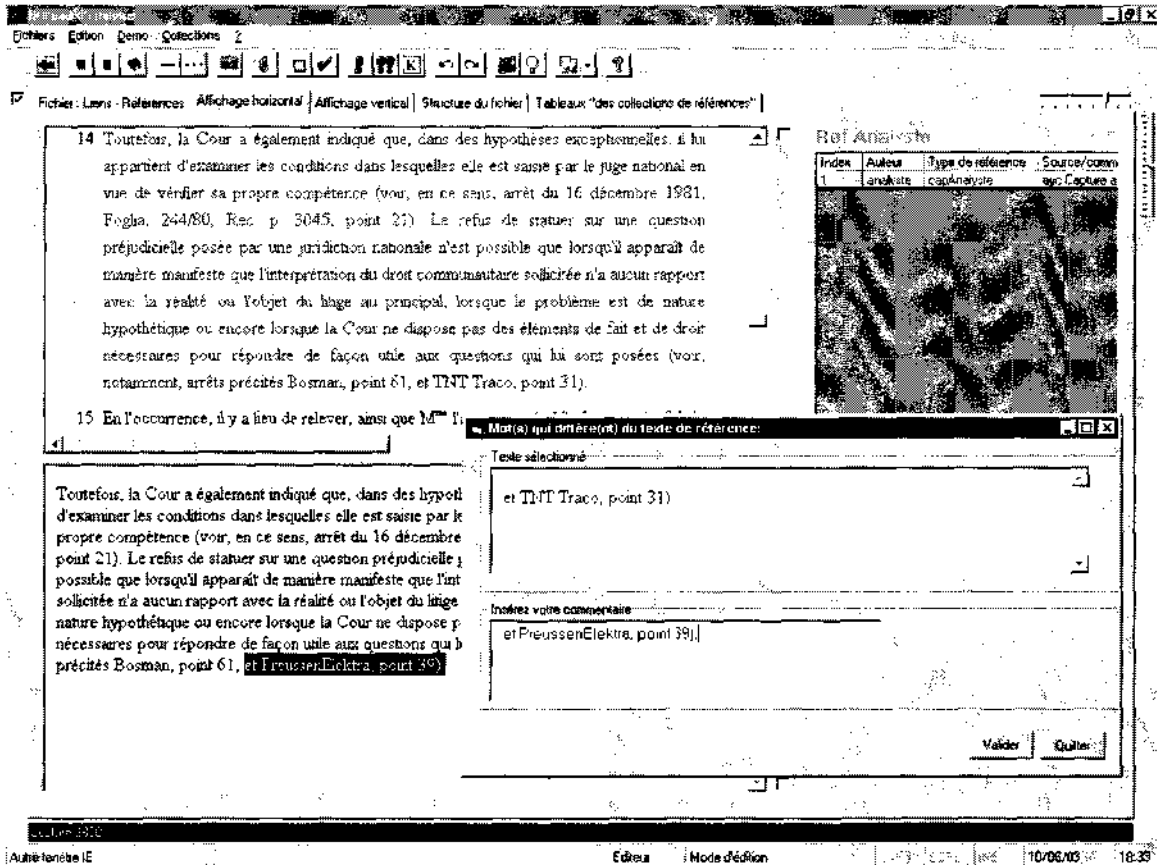


Figure 11: comparison of originator's final text with citation source

He can, for example, indicate that the reference of the Preussen Elektra case (source text) was replaced by a reference to the TNT Traco case (analysed text).



44. The background file grows richer as the commands inserted by the analyst are executed.

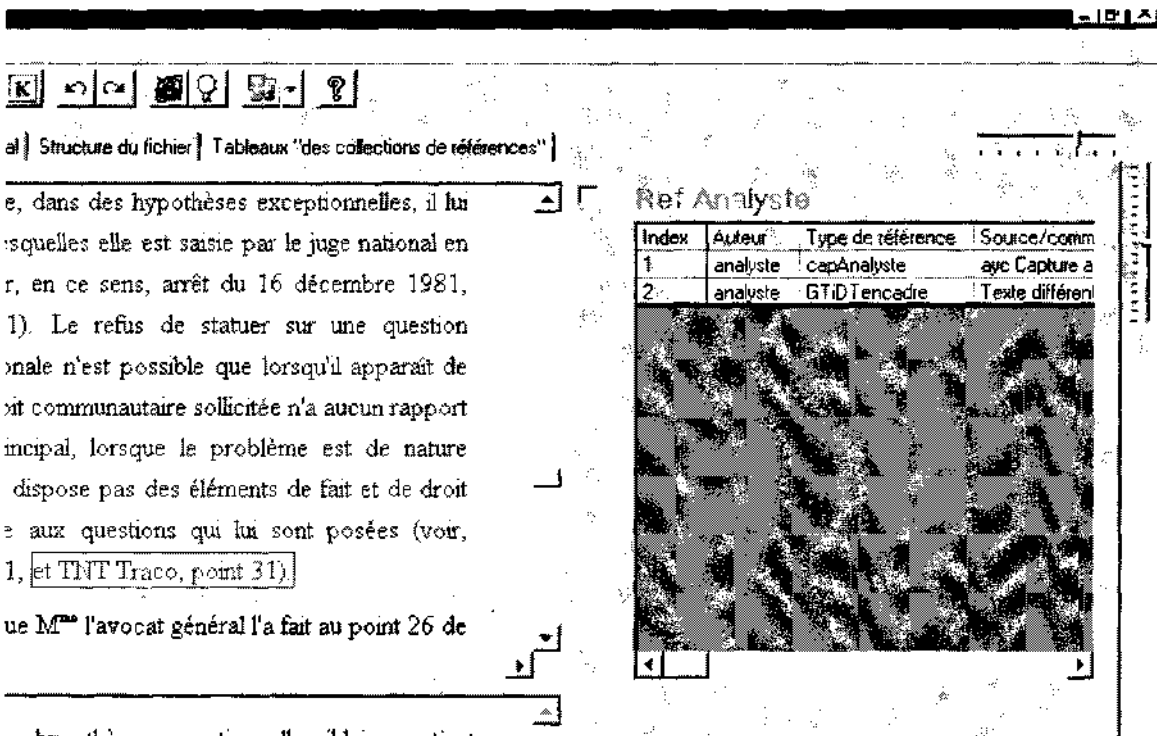


Figure 12: updating background file

2) Visualisation of the analysis

45. The analysed document contains a whole range of information relating to the text itself; the analyst need only run the mouse over the marked-up highlights to visualise them.

415/93, Rec. p. I-4921, point 59. et . point 38).

14 Toutefois, la Cour a également indiqué que, dans des hypothèses exceptionnelles, il lui appartient d'examiner les conditions dans lesquelles elle est saisie par le juge national en vue de vérifier sa propre compétence (voir, en ce sens, arrêt du 16 décembre 1981, Foglia, 244/80, Rec. p. 3045, point 21). Le refus de statuer sur une question préjudicielle posée par une juridiction nationale n'est possible que lorsqu'il apparaît de manière manifeste que l'interprétation du droit communautaire sollicitée n'a aucun rapport avec la réalité ou l'objet du litige au principal, lorsque le problème est de nature hypothétique ou encore lorsque la Cour ne dispose pas des éléments de fait et de droit nécessaires pour répondre de façon utile aux questions qui lui sont posées (voir, notamment, arrêts précités Bosman, point 61, et TNT Traco, point 31).

15 En l'Texte différent : mot(s) < et TNT Traco, point 31. > différent(s) de < et PreussenElektra, point 39). > M^{me} l'avocat général l'a fait au point 26 de ses conclusions, que la juridiction de renvoi précise qu'aucun grossiste n'est intervenu lors des ventes des produits pour lesquels Yorkshire a accepté des bons de réduction de la part des

Figure 13: analysed text - analyst comments displayed

46. When the analyst considers that the work has reached an appropriate stage, he can generate a translation template in any language of his or her choice and compare the analysed document with that outline. It is possible to call up a synchronised display of both documents, which the analyst can customise (vertical or horizontal presentation of the paired versions). This bilingual display allows him to check whether the textual references match the capture orders in the analysed document.

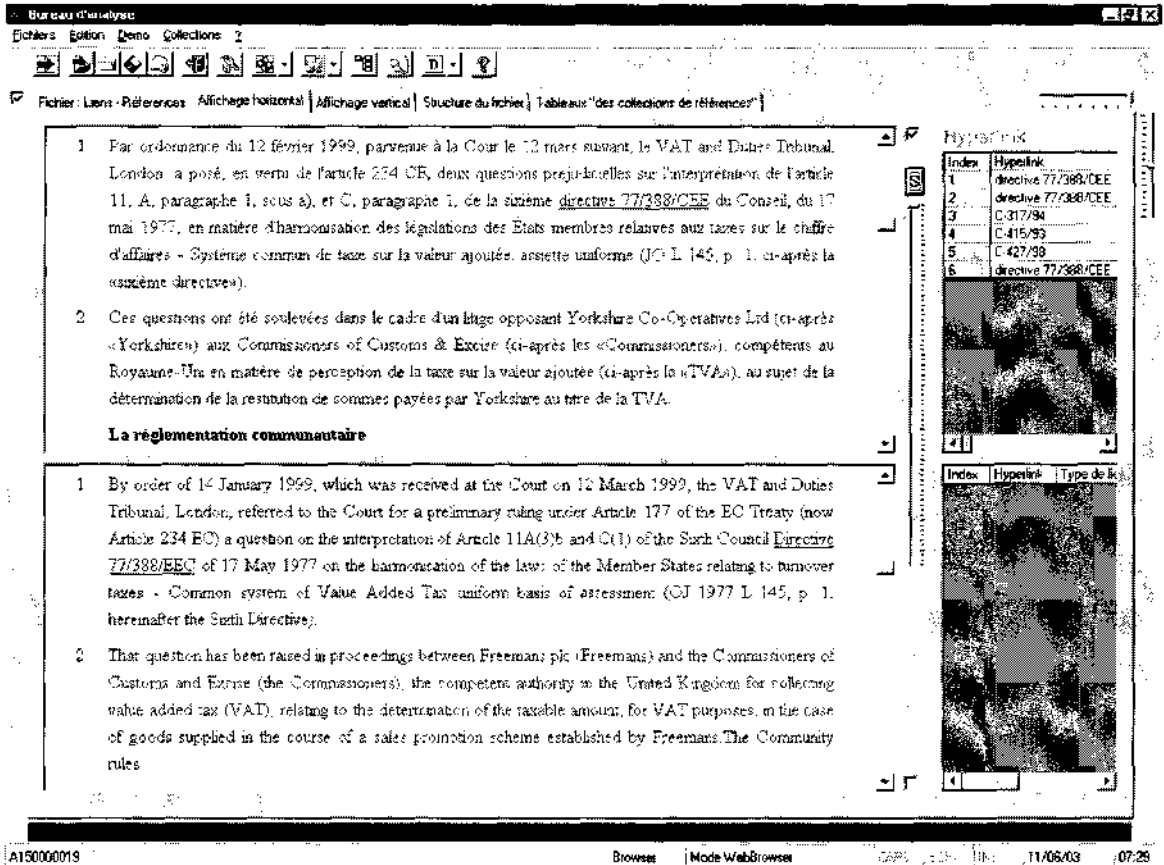


Figure 14: bilingual, synchronised display of analysed text with "translation template" in language of the case

3) Generating translation templates

47. When the analyst considers that his or her work is finished, he confirms the analysed document. A batch procedure checks in the "Suivi" database which target languages are required for the analysed document and produces a translation template in XML format for each of them. This outline is saved with the corresponding "background file" in the appropriate language in the document bank. Next, an e-mail is sent to each language division's local planning desk. The e-mail contains, in the form of hyperlinks, all the information to allow the team in charge of the translation of a document to retrieve from the document bank the translation template (automatically converted into Canvas format to receive the translation), the original document and the analysed document.

4) Translation

48. The translating environment is similar to the drafting environment.

49. The lawyer linguist is provided with the translation template in Canevas format. The outline contains the phrases and captures in his or her language; links give access to source documents which may be called up as a bilingual synchronised display. The parts of the texts which remain in the original language are treated as insertions made in “revision” mode. The lawyer linguist has two options: either to accept those parts of the text in the source language (in which case, the translation is typed over it to prevent a recapture of non-language dependent data) or to reject them (in which case, empty, pre-formatted, pre-numbered paragraphs are left in place).

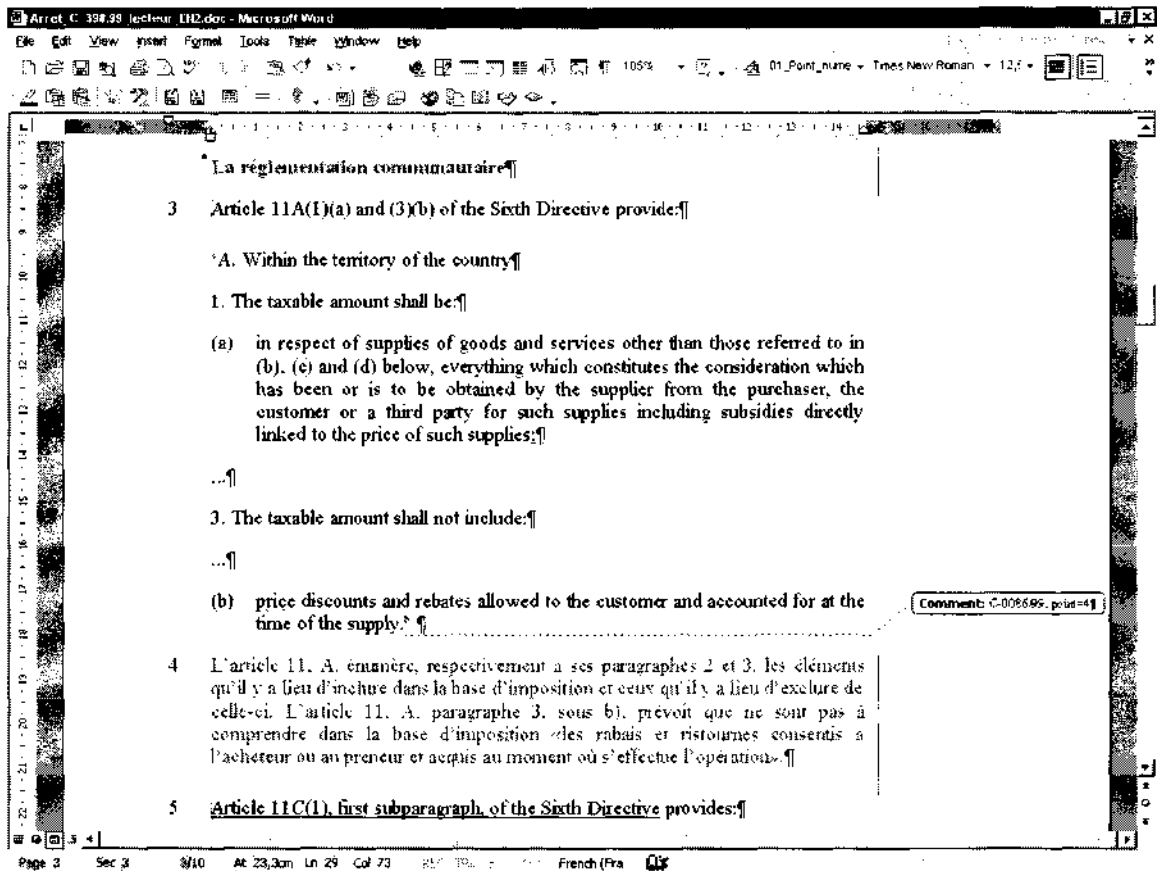


Figure 15: translation template in the wordprocessing environment

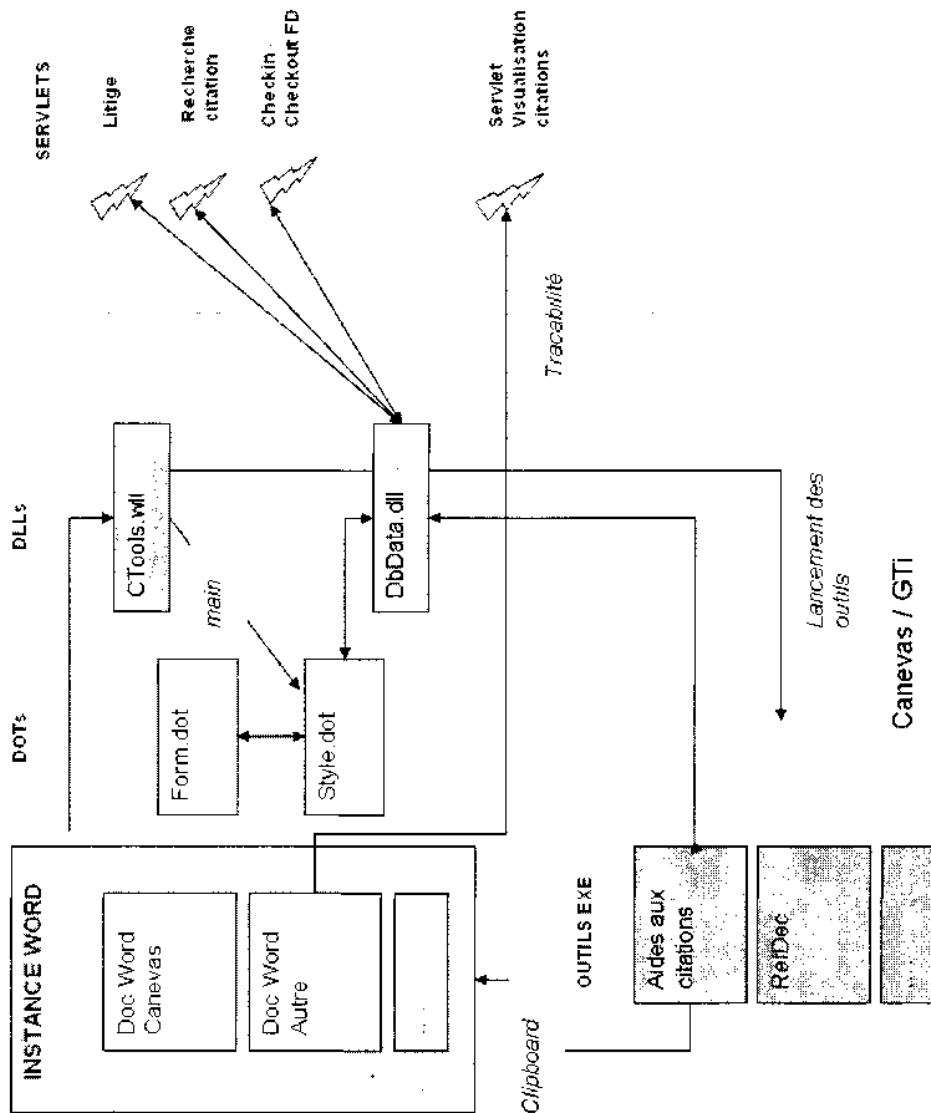
50. A number of tools allow the lawyer linguist to range over the original and the translation in synchronicity, print the bilingual version of the document (to re-read or to review it), compare successive versions of the same document or, finally, compare a part of the original text with its own source (used as the basis for the captures inserted in the translation templates). This last tool enables the lawyer linguist to identify the potential changes made by the originator of a text to the citation and to reproduce them faithfully in his or her own language.

5) Publication

51. The choice of XML as the central format for all documents for publication in the case law will considerably reduce the amount of processing required by documents for their electronic or conventional publication. Automatic processes will eventually make it possible to produce from such XML files not only HTML versions, rich in hyperlinks, of documents put on line on the Curia website, but also versions for uploading to Celex. Traffic between the Court and the Office for the Official Publications of the European Communities (OPOCE), which is responsible for the publication of the Court Reports, would also be reduced.

V – Putting the chosen approach into operation: technical architecture

Architecture technique globale



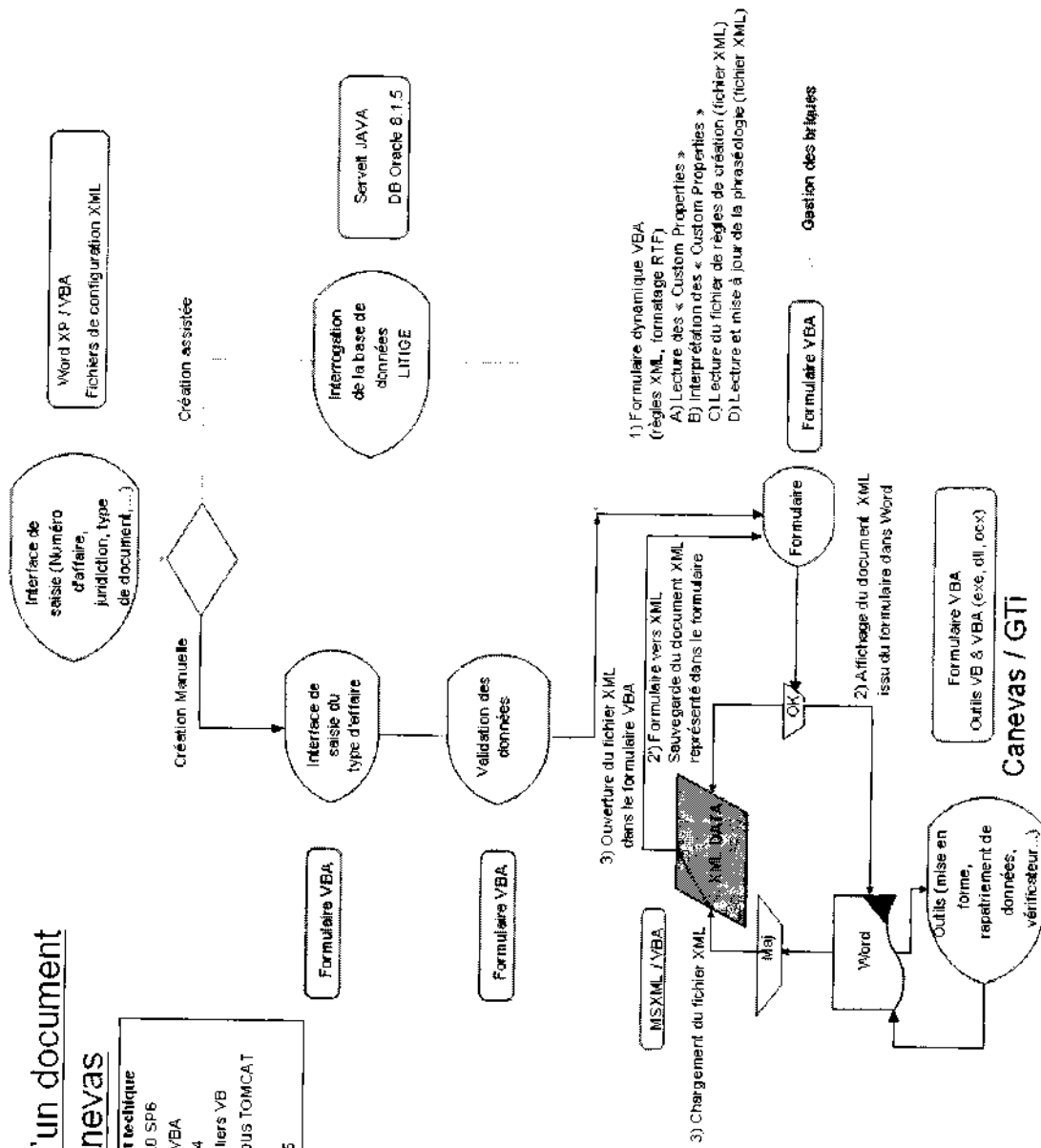
5 juin 2003



Création d'un document

Canevas

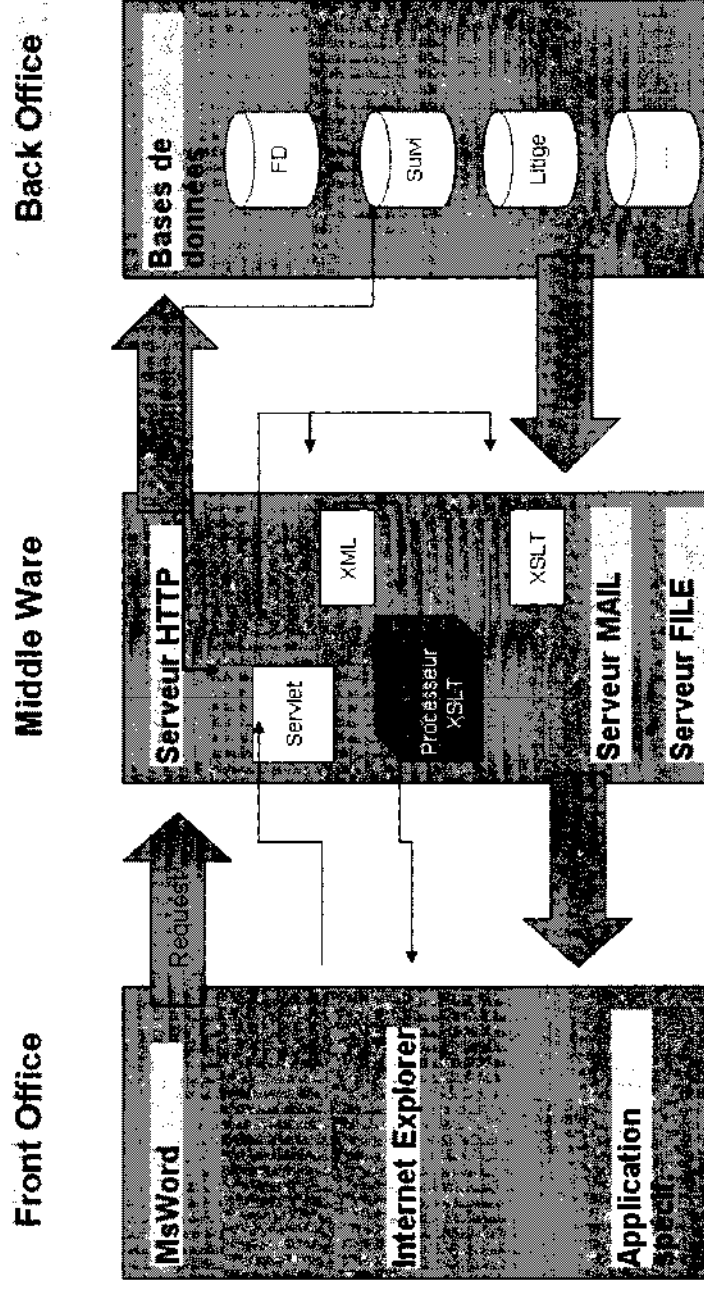
Environnement technique
 Windows NT 4.0 SP6
 Word XP / VBA
 MSXML 4
 Outils particuliers VB
 Servlet JAVA sous TOMCAT
 XSL
 Oracle 8.1.5



5 juin 2003



Prétraitement, analyse et
génération des matrices de traduction



5 juin 2003

Canevas / GTI



VI - Glossary

Celex: Acronym of the *Communitatis Europae Lex*. Celex is a set of parallel monolingual data produced by the institutions of the European Union and by the Office for the official publications of the European Communities (OPOCE). Celex provides full-text versions of all Community legislation (primary and secondary) and of the case law in all the official languages of the Union. Since 1997, Celex has been accessible on the internet (<http://europa.eu.int/celex>).

The translation Directorate: language service working to the Court of Justice and the Court of First Instance, responsible mainly for translation of the case law.

The translation Directorate employs more than 438 staff (as at 1 July 2003), of whom more than 300 are lawyer linguists, thirty revisers and management assistants and approximately 100 secretaries. The staff is divided into 11 language divisions a general service division and the “Organisation and Methods” department directly attached to the director.

All the language divisions except the French division do essentially the same work. For every new case, they translate the document initiating proceedings (Reference for a preliminary ruling) or the communication, drawn up in the relevant Registry (Court of Justice or Court of First Instance) summarising the subject-matter of the dispute and which is published in the Official Journal of the European Communities. The division whose language is also the language of procedure of the case has to translate a certain number of documents for the case file: observations of the Member States, applications by third parties for leave to intervene, the report of the hearing and the Opinion of the Advocate General for cases before the Court of Justice. In every case, they also translate the judgment, the summary and the communication of the judgment published in the Official Journal.

French is the language of the deliberations, so the French division has a special role. It always translates the document initiating proceedings and all the pleadings. In the cases before the Court of Justice, it also translates (like the other divisions) the Opinion of the Advocate General. It never translates judgments.

The general services division is responsible for tasks directly linked to translation which necessary for the whole Directorate: management, follow-up of the translations and the deadlines, analysis of the texts to be translated, pre-processing and production of translation templates, management of standard phrases, documentary and terminological research, management of archives, etc.

The department of Organisation and Methods is responsible for projects, not directly linked with the translation of cases, but which is essential to the smooth running of the Directorate: specific training sessions, participation in computing projects involving the Directorate, preparation for enlargement, coordination of working groups (secretariat, computing ...) or user pilot project groups (voice-recognition, intranet management, Canevas 2...), establishment of common working methods, management of the intranet of the Directorate.

In 2002, The Translation Directorate translated 395 000 pages corresponding in total to more than 35 000 request of translation. Nearly 60% of the pages were translated from French, the other languages being in variable proportions.

Document bank: System of electronic management of documents in course of being developed which will be put into production by stages from the last quarter of 2003. The document bank is designed to deal with all the documents produced or used by the Court in future: scanned pleadings, documents for publication in the case law, various internal documents, Community legislation (only texts referred to in Canevas documents and thus automatically repatriated to be exploited in the context of the multilingual production of documents), other external documents as required by the document bank (international conventions etc). The document bank is intended to cover archiving needs of the Institution the need for documents of the various users (multi-criteria and full-text searches, synchronised bilingual display, etc); it should also enable a dynamic follow-up of documents (management of the versions in various stages, rights of access, levels of confidentiality, etc) and conformity of the documents with the structural or descriptive rules adopted (XML format, metadata, etc) to be checked.

Case law: The entirety of the texts produced by the Institution in the exercise of its judicial work (judgments, opinions of the Advocate General, orders, opinions of the Court, summaries)

Lawyer linguist: Lawyer who completed his or her law study and has a thorough knowledge of at least 2 Community languages in addition to his or her mother tongue. The Court of Justice exclusively employs lawyer linguists (rather than translators) to translate the Community's case law (see the web site regarding recruitment and free lance work).

The official languages: initially 4 (German, French, Italian, Dutch). At present, 12 (Spanish, Danish, German, English, French, Greek, Italian, Dutch, Portuguese, Finnish, Swedish; Irish, which is an official language, is nevertheless not a working language). With the historic enlargement of the European Union on 1 May 2004, there will be a further 8 languages (Estonian, Hungarian, Lithuanian, Maltese, Polish, Slovak, Slovenian, and Czech).

Translation template: document automatically produced from various marks inserted in the document produced by the originator of the text using the Canevas application and by the lawyer linguist who has analysed the original document. The outline gives the lawyer linguist responsible for translation: the structure and lay-out of his or her document, the procedural data of the case (inserted in the restricted parts of the document), the standard phraseology in his or her language, the citations captured in his or her language and, finally, hyperlinks to the documents referred to in the original.

Litige: Procedural database set up and maintained by the Registries of the Court of Justice and the Court of First Instance and used to track proceedings.

Ellipsis: the part of the citation not quoted, shown by three points (...)

References to judgments: The reference to judgments is a particular kind of citation used in the case law in which the format for each language follows a different standard. Consultation of the references using Canevas tools permits references to be inserted in the document and ensures that the format will be correct.

Suivi: Database managed by the Translation Directorate to keep track of its activity (staff management, management of translation requests and the execution of the work, statistics, management of documentary or terminological information notes in relation to each case).

Computer-aided Translation (CAT): not the automated translation

(AT) used by SYSTRAN in the European Commission. CAT tools include software for management, terminology extraction, systems based on the physical alignment of multilingual documents and management of “translation memories”.

Vade-mecum: Compendium including rules for typing and citation, presentation of the first and last pages of judgments and the standard phrases to used in the various actions. The Vade-mecum is not translated but written in all languages in conformity with written style and typing practice for each language group or country. Both Courts have their own Vade-mecum, but they are generally referred to as if they were a single work.

XML: Extended Mark-up Language. Standardised language used to create structural documents where re-use and portability are guaranteed by compliance with the standard. The main advantage of XML mark-up is that it permits the logical alignment of the different language versions of the same document, rendering redundant the physical alignment carried out by translation-memory systems. HTML (Hyper Text Mark-up Language) is a language related to XML (but less complex). It is used to structure data to be accessed using Intranet/Internet browsers.