

Multilingual document management and workflow in the European institutions

Joanna Drugan

Centre for Translation Studies

University of Leeds

Leeds, UK

j.drugan@leeds.ac.uk

1. Background

2004 represents a pivotal moment in the evolution of language services in the institutions of the European Union. First, the simultaneous accession of ten new member states and their corresponding languages has meant a huge increase in translation workload and potential language pairs literally overnight. Second, and of particular interest to anyone working with electronic translation tools, the more efficient integration of a range of translation software and tools to manage multilingual projects is now under discussion at inter-institutional level, via the newly-created working group on tools and workflow. In its quest to identify the most efficient working methods and appropriate integration of translation tools, the group can draw on nearly a decade's use by EU translators of the Trados Translator's Workbench (TWB) translation memory (TM) tool, which has increasingly been used in conjunction with a package of other translation support tools including machine translation, in-house workflow management tools and voice recognition and dictation tools. The establishment of the working group is also timely given the obligation this year for the EU to launch a new Call for Tenders for the supply of its TM tool. Trados's TWB and MultiTerm terminology software were originally designed in conjunction with the institutions to address the needs of in-house staff but the public body's requirement to ensure value for money means it is now obliged to

consider rival tools, most of which have evolved significantly since the advent of XML and Unicode.

Moreover, since TWB was made available to all European Commission translators in 1997ⁱ, translators and former terminologists have built up substantial resources in electronic formats and have developed a range of working methods in the different language units, or indeed, different cities in which they are based. How can these resources be used and managed as efficiently as possible to address the increased translation needs of the EU, and how can computer tools help with this task?

This paper will address these questions by presenting original findings from a recent six-week research placement during which the author was based in the English Unit of the Directorate General for Translation of the European Commission, with additional access to translation tools specialists, members of the committee on translation workflow, translators from both pre- and post-accession member states, Heads of Units and staff in other institutions, including the Parliament, Court of Auditors and Translation Centre, both in Luxembourg and Brussels. Research took the form of tailored interviews, work shadowing of on-site and teleworking translators, performing translation tasks using in-house tools and observing training in translation tools. Issues relating to translation workflow and data management were thus investigated across different languages and institutions, resulting in a comprehensive comparative overview of effective translation management strategies within an expanding multilingual non-profit organisation.

Why might this be of value for others in the translation and computer communities? As Catriona Picken pointed out to a much earlier Aslib readership, following the development of two early EU computer projects AVIMA and GEPROⁱⁱ, 'It can be no coincidence that large organisations with an enormous translation turnover are giving the lead here. They have both the motivation and the resources to carry through projects of this kind, and other, smaller organisations with more limited resources can learn from these experiences what can be done in this area' (1991: vii-viii). By summarising key computing and translation problems identified by the EU institutions and assessing their attempts to address these, a new way of looking at our own practice as translators, tools developers or trainers is made possible. In addition, the EU institutions employ a wealth of highly-skilled practitioners who represent an unusual resource in the translation sector

in that they generally work in the same building in very large groups of multilingual colleagues. Communication among such colleagues is likely to have helped in the early pinpointing of problems relating to particular language pairs or ways of working, for example. Finally, like any public body, the EU is obliged to publish and justify expenditure and this has meant its staff adopting a practical problem-solving approach to the problems they have identified. The EU has not been a forum for research-led investigations of translation theories but a voracious consumer of ever-increasing quantities of translation to tight deadlines, with a resulting ambition to use electronic tools efficiently. This generally means the institutions have produced concrete solutions which can be outlined for consideration by others facing similar problems.

2. Multilingual document management, workflow and technology in the European institutions

Before outlining the translation problems which EU institutions have identified and then addressed using computer tools, it is important to have a clear picture of typical translation workflow for in-house translationsⁱⁱⁱ. This is less standard or straightforward than might be imagined. First, EU translators work for an unusually diverse assortment of translation requesters, from a newly-elected MEP using translation for the first time to the members of long-standing committees which would be unable to work without constant recourse to translation and whose members therefore become sophisticated, demanding consumers. Deadlines and the balance of a translator's workload are accordingly varied and flexible. Second, the type of document^{iv} sent for translation and its intended readership are similarly varied. Translators might work on a short handwritten letter from a citizen of a member state or on the nineteenth draft in electronic format of a confidential amendment to controversial legislation. Third, historical, political and geographical factors complicate matters, with individual institutions like the Parliament and Court of Auditors having their own internal translation units which can approach translation differently than those in the longer-established DG for Translation in the European Commission. In addition to the range of institutions with internal translation services, smaller EU organisations and units can call on translators and tools specialists based in the recently-established Translation Centre, a unit independent of the DGT which provides translation services for EU bodies which do not have their own internal translation provision. Established at different times to meet different needs in a different number of languages, it is hardly surprising that all these translation units evolved their own working methods and workflow.

Nonetheless, certain key stages are now common to most of the translation providers and can be outlined below. Sources of information used to compile this outline of typical workflow include official publications at www.europa.eu.int and interviews with translators in various EU institutions and with members of the Brussels-based Pre-Processing Unit of the DGT for Translation at the European Commission. Where common exceptions or several options exist, they are noted. It should also be stressed here that one of the aims of the new working group on workflow is to identify and share best practice across the institutions, so a more standardised approach is possible in the near future.

1. The process begins simply enough with a **translation request**. This comes from a DG or Service and contains standard information, notably the target language(s) into which the source text (ST) is to be translated and the deadline by which it is required. The ST is attached, though it may not be in electronic format.
2. Most European institutions have some form of **Central Planning Unit** which accepts and logs the translation task.
3. With some exceptions (e.g. the European Parliament, the Court of Auditors), most tasks are sent to the Commission's **Pre-Processing Unit**. This Unit automatically:
 - a. generates a *Fiche de travail* for each language pair into which the document will be translated. This is a cover sheet indicating the job number, deadline, commissioner of the translation and any useful material recommended for reference^v. Where possible, a copy of the electronic ST file is generated for each language pair in the in-house electronic Dossier Manager interface and any relevant notes are inserted;
 - b. performs an initial check for similar texts previously translated and indicates their codes as reference documents on the *Fiche de travail*;
 - c. runs a check on the central TM and the Celex legal database then indicates any useful material on the *Fiche de travail* and in Dossier Manager;
 - d. decides whether to recommend using TWB to translate the document via a check box on the *Fiche de travail*;

- e. sends a dossier containing hard copy and the *Fiche de travail* to the Head of Unit concerned.
4. The Head of Unit receives the translation request and assigns work to a named translator or places it for collection at a central area as appropriate.
5. The translator collects the dossier and creates the **Target Text** (TT) file for her language pair in Dossier Manager.
6. The translator, or sometimes a secretary, locates suitable reference material from the **Euramis**^{vi} central server (with some exceptions). If previous translations might be useful but are not currently aligned and stored in the TM, a request is made by the translator or a secretary for the material to be aligned.

At the stage when actual translation begins, approaches diverge according to the ST format and to translators' preferred working methods and skills.

Table 1: 'Typical' translation workflow

Translation using TWB	Translation using Word	Translation using dictation
7. The translator opens the TT file in Word and uses a pre-installed TWBTools macro to create a new TM with one click. A dialog box displays all relevant TM material to be imported into the new working memory.	7. The translator may work from the original ST sent with the translation dossier if no electronic copy exists. Otherwise, the translator opens the TT file in Word where colour-coded suggestions from reference material are displayed. For example, 100% matches from a TM will be inserted in the TT but coloured differently than fuzzy matches.	7. The translator may work from the original ST sent with the translation dossier or read from the pre-processed file as for translation using Word.
8. The translator clicks on the Open/Get button in the standard TWB toolbar to start translation, with 100% or fuzzy matches colour-coded in the TT segments.	8. The translator types over the ST, accepting or editing suggestions from the TM as he proceeds through the document.	8. The translator dictates her TT into a tape recorder.
9. Translation is completed and saved. A printed copy is usually sent for revision by a colleague.	9. Translation is completed and saved. A printed copy is usually sent for revision by a colleague.	9. If the translator uses Dragon Naturally Speaking, she post-edits the dictated text. If not, a secretary types the translation for checking by the translator.
10. After revisions are made, the translator uses a TWBTools macro to Cleanup the TM.	10. Revisions are made.	10. Revised translation is completed and saved. A printed copy is usually sent for revision by a colleague.

11. The translator sends the working TM to the appropriate thematic sector of the central TM via the Euramis server.		11. Revisions are made.
--	--	-------------------------

At this point, the translating unit releases the translation and it is archived electronically in Euramis so that it can be used by future translators working on similar texts. For example, a translation which has been dictated can later be post-aligned for a future job taken on by a TWB-proficient translator. Finally, the translation is sent electronically to the requesting department and statistics are compiled on the translators' use of tools, number of pages translated and so forth.

In moving through these stages of typical workflow, EU translators all have access to the CELEX database of legal texts and electronic terminology lookup. Most translators (all in many institutions) also have access to TWB and an increasing number are choosing to invest time in the initial preparatory work needed to use Dragon Naturally Speaking voice recognition software, though this is not available as standard^{vii}. Translators in a limited number of language pairs^{viii} have access to Machine Translation(MT) and some can still call on secretaries to audio type their dictated text.

Having outlined the range of computer tools available to in-house translators and their diverse working methods, we can now ask what problems have been identified in recent years by the world's largest supplier of multilingual translation. In the past, the EU translation services have acted as a motor driving forward technological advances so it is not surprising that they have focused on efficient use of translation tools in order to address these problems. We will also note how far such solutions have been implemented within and across the EU institutions.

The recent growth in demand for translation and shortage of suitably-qualified language professionals are problems for the industry as a whole, but they are undeniably more acute for the EU institutions in the context of enlargement. The related growth in staff numbers, range of languages and data produced and stored has created and exacerbated certain problems.

First, the growing demand for translation within the institutions will certainly not be matched by a parallel increase in in-house staff numbers. Over the past five years (i.e.

before enlargement), the average annual increase in demand for translation at the Commission alone was 5.3%^{ix}. This has two implications: efficiency must be maximised among in-house translators; and more translation must be done in other ways than the traditional method of using this pool of in-house staff. How has the EU addressed these two points so far?

The need for greater speed in the production of translations without a corresponding drop in quality has been addressed in part by identifying tasks which can be done by other staff. A happy coincidence here has been the increasingly well-developed word-processing skills among translators, meaning that secretaries who only recently spent their working days audio- or copy-typing have more time to devote to electronic translation support tasks. The EU institutions are fortunate to have as a resource teams of linguistically-skilled secretaries^x who can thus be trained to carry out initial preparatory tasks such as retrieving previously-translated material for translation jobs. A simplified in-house alignment tool, based on Trados WinAlign but with less preparatory input needed from the end user, means secretaries can also post-align previous translations and their STs so that a greater quantity of TM segments is available for future translation matches, again increasing translators' future efficiency. In this way, growing demand for translation can even be presented as a virtue, with more varied workloads for support staff, more extensive use of translation tools and a greater likelihood for translators of finding matches in larger TMs.

Behind the scenes, more staff have also been allocated to such tasks as running the TWB helpdesk, which provides instant technical support via phone or email. In interviews, translators repeatedly made unprompted positive comments on the valuable resource such help represents, with some commenting that they were more likely to use TWB for translations (hence increasing their speed and efficiency) precisely because they knew help would be available should anything go wrong.

Encouraging staff to make full use of the tools available is clearly one important way to increase efficiency, with the obvious qualification that those who are uncomfortable using the tools might actually become less efficient or even choose to leave. Perhaps for this reason, the EU institutions favour a voluntary approach. Translators are encouraged to adopt the tools by subtle means such as a rolling voluntary one-day training

programme in TWB and Euramis use. This is possible because the Commission's DGT has automated many of the cumbersome stages involved in using TWB, notably through in-house macros and toolbars in the Word interface. Staff can thus easily become adequately proficient users in one working day. Importantly, a recent decision has been taken to provide training not by tools experts but by practising in-house translators who themselves choose to use the tools and who volunteer to spend a day or so per month training others in teams of two. Their evident enthusiasm and ability to relate the training to actual in-house translation scenarios make the training highly focused and effective.

One exception to the voluntary training principle has been the group of translators who have arrived *en masse* following recruitment to the accession countries' languages. An additional programme of training in TWB and Euramis use was put in place and all those who passed the first competitions for the new languages were given a place. Given the huge backlog of multilingual documents for translation by these new colleagues, however, they were evidently keen to find out how they might eliminate repetition in their heavy workloads and in interviews seemed unanimously to welcome the opportunity for training in TM.

Another way the EU has encouraged staff to work more efficiently is through the simple step of making a wide range of tools available. If translators can choose to use dictation and voice recognition software or MT, more are likely to find one tool which suits them. Even those who prefer to use only MS Word are encouraged to exploit the time savings offered by translation tools through the automatic insertion of previously-translated segments in the TT at the pre-processing stage. With some revisions of documents offering over 90% of the text as exact matches, translators see for themselves why it makes sense to work this way for at least some jobs.

Greater efficiency has not only been addressed at the level of the translation units' staff. An effort is being made to educate those requesting translations and to control document production. For example, the English unit of the DGT for translation requests that only electronic texts are sent for translation and some highly repetitive translations such as Calls for Tender have long been mostly automated. Here the size and diversity of the institutions are obvious stumbling blocks to reform, however. The backlog of translation had reached 60,000 pages by 2004 and predictions of 300,000 pages by

2007 pushed the Commission to institute a policy of demand-management limiting requesters to shorter document lengths^{xi}.

Nor can political factors be ignored in any consideration of efficiency in multilingual document production and translation. For example, one other potential way in which the growth in demand for translation could be addressed might be to decide not to translate certain documents, or to restrict translation to a core group of languages. STs might also be authored in a restricted group of languages. However, the EU understanding of the requirements of democracy^{xii} means that all languages need to be available for at least some documents. Even here, however, there has been a drive to greater efficiency, with the principle of two or three working languages for committee discussions now accepted. Documents resulting from the discussions can then be translated into all official languages or a wider sub-group of languages only where this is appropriate. Official statistics on the languages in which Commission STs are produced make this development clear.

Table 2: Evolution in source language of texts translated by DGT translators

Year	Total no. of pages	SL English (% of total output)	SL French (% of total output)	SL German (% of total output)	SL other EU languages (% of total output)
1992	914 649	35.1	46.9	6.2	8.8
1997	1 125 709	45.4	40.4	5.4	8.7
2002	1 302 313	57.4	29.1	4.6	8.8
2003	1 416 817	58.9	28.1	3.8	8.9

A further way to reduce translation requests has been to make MT available to non-translators across the institutions. Statistics show an impressive increase in MT use for those language pairs in which it is available, with many staff happy to use quickly-available MT output to get the gist of an email, for example, rather than sending a translation request.

Table 3: Total use of MT in the European Commission

Year	Total no. of pages
------	--------------------

	produced
2000	546 000
2003	875 584

Increasing volumes of translation only serve to highlight underlying issues of multilingual document management in the institutions. In many ways, the EU is an extreme example of the challenges facing any multinational organisation. Documents are rarely completed in one language never to be revised. More commonly, discussions will take place in two or three languages in a committee whose members are frequently not native speakers of any of these languages. Resulting proposals will be sent for translation into a wider range of languages (up to 19 since enlargement), at which point they might be discussed via interpreters in Parliament with questions and answers in any official language. There may be public consultation on revised proposals before further debate, amendments and final drafts before the legislation is translated into all official languages for application in member states. It is unlikely that the same translator will be available to follow the entire process through, and entirely probable that similar texts or amendments will need to be translated in future.

Such working methods raise problems of consistency, document version management and efficient data storage, many of which are amenable to technological solutions. First, requiring all translators to use standard official terminology sources is essential. As documents are discussed in committees, Parliament and in written form, consistency across institutions has also been necessary – and problematic. Here, the institutions have recently acknowledged the need to work together and a new development, IATE (Inter-Agency Terminology Exchange) will incorporate the terminology databases formerly accessed via Eurodicautom in addition to other institutions' or individual units' terminology databases and alignment materials. Sharing terminology and ensuring that translators access it via one interface has clear advantages in addressing the problem of consistency and updates. The database can be structured to make the most appropriate terminology suggestions first - for example, if a term is found in a termbase of official legislative terminology, that will take precedence over a different match found in an individual language unit's subject glossary, though all matching terms will be displayed with their context and source for the translator to make the final decision.

If one solution is efficient sharing of official terminology, using technology to manage version control is also key. This is typically managed at the TM or pre-processing level. The first suggestion made to the translator for a segment found in an existing TM is always the most recently-entered matching segment, with previous versions available for reference at the translator's request and listed in descending date order. Notes features can be used to indicate any particular translation requirements across all the target languages into which a ST is to be translated.

Linked to the growing demand for translation, the growth in electronic data (size of TMs and terminology databases) represents a clear challenge for many in the industry. With more than 70,000,000 segments, 130 gigabytes of data and 10% growth per year,^{xiii} the European Commission's TM database alone might be likely to pose similar problems. However, this problem seems to have been addressed satisfactorily via certain precautions in TM design and management, particularly the decision to work with one central TM repository, Euramis, which can be managed efficiently by suitably-trained users. The DGT for Translation has identified key steps in effective data management:

- A thematic separation of TMs has been adopted (via 'clusters' such as policies or services);
- Translators may attach up to three TMs to any one translation project;
- Local duplication of TMs is avoided by using the central TM in Euramis for storage;
- Translators cannot upload the same material to the central TM twice, avoiding duplication of entries;
- Only a small sub-group of trained expert users can reorganise or maintain the central TM;
- Translators' suggestions for modifications to TMs or terminology databases are sent to central contacts for checking before amendments are made.

Multilingual document management has been further complicated by the increase in numbers of staff and the range of languages into which they translate following enlargement. Again, consistency represents a central problem, particularly since translators working into new EU official languages will be building up the core terminology and TM resources their colleagues will rely on in decades to come. The EU institutions have responded to this problem and to that of empty TMs for the new

languages by tailored early training in use of translation tools. New translators are encouraged by frequent statistical feedback to use electronic tools as much as possible: they can observe how the terminology and TM segments are growing thanks their input, often in quite small units of only a handful of staff. In addition, support staff have aligned the initial *acquis* of legislative texts already adopted in each member state's national parliament as an original TM. Another way to address the sudden rise in the number of official language pairs and concomitant paucity of resources for some languages has been to make indirect translations available via Euramis. For example, TM segments from a document which has been translated from French to English and from French to Latvian can be associated to form an English-Latvian language pair.

The Euramis project aims to address other problems relating to enlargement and multilingual document management too, sometimes in conjunction with another tool, Dossier Manager. In an expanding multilingual organisation with huge quantities of documentation, staff are faced with very basic questions. How does a translator know a text hasn't already been translated or that another colleague hasn't also just started the same job? How can the most effective use be made of existing translations, to avoid both duplication of work and (unintended) multiple translations for the same ST? When a ST is to be translated into multiple target languages, how can queries about sense of the ST best be managed, so 19 translators don't have to ask the same question of the requester, for example? How is a translator to decide which TMs will be most useful for a particular translation task?

The Euramis interface aims ultimately to address all these problems via a 'one-stop shop' for translators, providing integrated access to a range of features so that resources are shared as widely as possible and translation tasks are tracked and managed effectively. It currently combines with another tool, Dossier Manager, for work tracking purposes in the Commission, but it is hoped ultimately to integrate Dossier Manager's functions with Euramis to create a new Translation Workstation. Dossier Manager is the interface in which translators in the Commission take on new jobs. Only one user can register for a new job, avoiding basic duplication of work. A note feature allows communication between translators working from the same ST into different target languages. A recent development has been the move to using a 'lead translator' for each such multilingual task. This translator, usually the first to start work and on a voluntary

basis, will liaise with the requester then make suggestions and clarifications available to all working on the project via the note feature.

An important technology-based response to the problems noted above has been the increased sharing of resources and knowledge at inter-institutional level. Euramis began as a European Commission DGT for Translation project but is to be enhanced and made available to those working in other institutions too. Features supplied by Euramis include:

- Automatic analysis of ST and retrieval of related previously-translated material;
- Access to central TMs (remote access to shared data);
- Access to MT for the language pairs currently supported;
- Access to the Celex legal database;
- Online concordance searches;
- Online consultation (e.g. document searches);
- Output of searches in TWB or Word format (user decides);
- Prevention of re-importing to the central TMs by users;
- Alignment of previously-translated documents;
- Checking feature to ascertain whether a document has previously been aligned by another user;
- Creation of reverse and indirect TMs from previously-translated documents;
- Improved maintenance of data via central management;
- Filter searches on attributes such as DG codes.

One growing group of EU translators who do not benefit from all the facilities available in Euramis are freelancers. As well as increasing internal efficiency to address the increasing workload, the EU institutions have recently accepted the need for a higher proportion of translation work done by outside contractors. New targets mean that 40% of all EU translation should be outsourced by the end of 2004 and 50% by 2007; currently the EU pays 30 million euros to translation contractors per year but by 2005, at least 80 million euros should be devoted to external translation^{xiv}. The institutions are thus revisiting the question of effective support: what materials should freelancers be given so that their translations can attain the same levels of consistency and quality as in-house work? Currently, the EU states that freelance translators are provided with 'useful background documents; the name of a contact who can assist with translation

problems; access to the translators' tools provided by the DGT, such as the terminology help service, the documentation centres, Eurodicautom and the CELEX database; feedback on the assessment of their work' (*Translating for a Multilingual Community*: 9). However, there is no access to the central TMs beyond segments inserted in pre-processed documents, nor to the various features available in Euramis (for example, alignment of previous translations), although some internal discussions have raised this issue in the context of the increasing proportion of outsourced translations.

This problem brings us to the broader issue of integration of translation tools. The increasing proportion of work done externally highlights internal problems too. Cost and access to tools are problematic for freelancers but also have implications for the EU institutions with their requirement as a public body to obtain value for money. Not all freelance translators have access to tools such as MT or TMs, but not all in-house translators in all institutions have access to these either. Various developments help address this problem. First, automated pre-processing of STs in Euramis means that material stored in the central TM can be made available without access to the TM itself, as 100% and fuzzy matches can be inserted in the target text before sending to the translator, whether he is working in-house without tools or externally. This facility represents a real advantage for translators, by cutting down on research time and improving consistency, and importantly, without affecting the rate freelancers are paid for the work. Second, offering support by phone and email via a helpdesk addresses technical problems quickly and, importantly for home-based translators^{xv}, cheaply. Finally, if a decision to share TMs is taken in future, developments in the industry will facilitate the exchange of files. The widespread introduction of TMX exchange format means freelance translators will be able to access and return TMs while using other translation tools, perhaps cheaper or free tools such as Wordfast.

Much of this relies on appropriately trained and willing staff and this has been another problem faced by the institutions, as by any large employer faced with dramatic changes in tools or working methods. Many staff with established working patterns will demonstrate reluctance to use tools effectively, even at all, perhaps through technophobia but also perhaps through self-interest. It is noticeable in Table 1 that those translators using TWB or Dragon Naturally Speaking have more stages to go through to complete a translation task, and initial training does still demand translators' time when

their efficiency is measured in pages of translation completed per annum. Of course, overall long-term efficiency gains in translation should outweigh the brief amount of time spent in training, creating new TMs or uploading finished projects to Euramis, but this will not be immediately apparent to all. In addition, translators for the new official languages face a dearth of prior terminological or TM resources and hence an absence of an existing incentive to adopt the tools.

The EU solution here has been two-fold: make it as easy as possible on the translator and lead by example rather than coercing staff. To make the tools as user-friendly as possible, the translation tools team have automated as many stages as practicable. Two notable examples:

- Instead of working through the various stages in the TWB Create New Translation Memory wizard, where they would have to choose language pairs, select termbases and TMs manually, enter attributes and so on, an EU translator simply opens Word and selects Create TWB Memory from the in-house TWBTools menu. The whole process takes less than a second and clearly eliminates both repetitive work for the translator and the possibility of human error corrupting the data in the central TM.
- Instead of analysing each new ST and searching manually for useful material in the TM, the pre-processing unit handles this automatically, attaching the relevant TMs for TWB users or inserting colour-coded matches in the Word document for others.

Many translation or document management tasks have also been taken off the hands of translators. For example, only a small sub-group of specially-trained translators is involved in TM management. This both lightens the workload of their colleagues and guarantees only the most appropriate content is held in the TM. Finally, an ongoing effort is being made to educate translation requesters so that using the tools is straightforward. Obviously, if a long document is received by fax for translation by a short deadline, it is unlikely to be translated using electronic tools. Simple factors such as requiring standard date formats or approved terminology for all STs are not near being resolved thus far, and arguably are unlikely to be, given the working context of MEPs and Commissioners being replaced at regular intervals. Training, inter-institutional communication and negotiation are likely to be necessary qualities for the foreseeable future.

The second approach has been to lead by example rather than coercing staff to adopt translation tools. Evidently, by inserting matches in pre-processed documents for all translators, even those not using TWB can see why it makes sense to store matched segments in a TM as they too benefit from the retrieval of previous work. By providing targeted training on a rolling basis, those translators who are interested to learn how the tools work are then given the opportunity to do so on a voluntary basis, making for a more motivated learning environment. These translators typically return to their units and are very satisfied to see fairly dramatic increases in their productivity, measured in pages per annum for each translator. Colleagues will observe this and be able to ask the newly-trained translators in their unit for advice and encouragement. Even simply making some tools available without additional training can help here. Thus 34% of the total use of MT across the European Commission in 2000 came from translation staff (*Translating for a Multilingual Community*, 2002 ed.: 11) and kept growing to reach 40% of total use by 2003 (*Ibid.*, 2004 ed.: 12).

Another problem in integrating translation technology has inevitably been human error or misuse of the tools. How can human fallibility be limited? How can staff learn to use tools as effectively as possible and to adopt new attitudes, such as storing work centrally so it can be shared rather than duplicating effort via local termbases and so on? Again, the EU, and particularly the Commission, have been moving towards greater automation of tasks to get round these problems. By fixing project parameters and settings centrally and prompting users for information, rather than encouraging them to design their own termbases or memories, irrational or inconsistent project set-up can be avoided and data will be compatible with other information already stored in the system. Training and clear handouts for later reference also address some common mistakes, such as cleaning up a translation and sending it to the central TM before the revision stage. Effective management and storage of data once it leaves the translator resolve some earlier errors. Attempts to share best practice across units or institutions can also help. The DG for Translation website, to which all translators log on each working day, provides a convenient, accessible forum for such communication, for example.

Inappropriate use of translation tools can also have an impact on quality in terms of translation content. One obvious problem is to accept matches suggested from the TM

when they are unsuitable for the particular context of the translation. To some extent, electronic tools are no different than paper dictionaries in this respect: a harassed translator might just as easily make inappropriate choices with more traditional tools. The Commission addresses such problems by flagging them up to new users during training, and by internal revision of translations for all new colleagues' work and for a sample of established translators' output.

Interference between different tools has sometimes posed problems for their integration in the EU. One frequently-cited example is interference between the in-house Celex legal database and TWB. This means that translators are often unable to complete the clean-up process. The typical solution here has been to turn to the Helpdesk for solutions and where necessary, adopt a pragmatic approach by post-aligning the work, for instance. Finally, there is a reporting mechanism whereby translators can feed back to the Translation Tools unit when similar technical problems arise so they can be resolved at the appropriate level if necessary.

The third broad area where translation problems are being addressed through technology in the EU institutions is that of workflow. These problems have been officially recognised recently by the establishment of an inter-institutional working group focussing on the issues, the aim being to identify and share best practice.

The size of the workforce is clearly an important problem for a group aiming to introduce even basic common standards in workflow across the institutions. The Commission's DGT for Translation was already the largest translation service in the world even before enlargement^{xvi}. Each individual, unit, language and institution has established ways of working and while there is clearly a great deal of common ground or shared practice, as seen in Table 1, discrepancies do remain. Any attempt to standardise elements of workflow will need to take this on board. As with the introduction of translation tools, the EU approach thus far has been gradual and based on encouragement rather than coercion. Some core stages are now established across the institutions, often by removing tasks from unit-level management so they are handled centrally, such as the pre-processing stage. Use of tools such as Dossier Manager to share information and track workflow is also achieving a degree of standardisation in working practice. Finally,

standard training for new colleagues and a mentoring system within units ensure that those joining the institutions learn appropriate workflow early in their careers.

A related problem affecting workflow is the diverse nature of the EU translation staff. First, there is of course diversity in the use of translation tools. One standard workflow model will not be suited to all translators, given the range of tools available or chosen by individuals. Second, geographical distribution of staff might pose problems for standard workflow models. In-house staff and services are spread over at least two sites, or three in the case of the Parliament, and all institutions permit home working for translators. The rising proportion of freelance workers in the near future will further complicate the picture. The chosen solution thus far has been to centralise and standardise preparatory tasks and post-translation tasks, but this may be less feasible as the proportion of outsourced work increases. Will it still be possible or desirable for secretaries to post-align translations and STs when 50% of documents are translated externally, often using TM tools anyway?

Enlargement has also changed the context in which translators are working in an important and controversial way: the balance of languages spoken in the EU institutions has shifted significantly (see Table 2). As the proportion of English STs increases and committee discussions take place in English more frequently, there may be pressure to use English as a 'pivot' language with implications for both workflow and staffing. For instance, if more translation is to be done out of English, often from STs written by non-native speakers, the editing service might usefully be introduced as a formal stage in typical workflow. Such solutions are likely to be challenging for many to accept, given the political pressure from some states to safeguard the historical role of their languages. Technology has a role here, as the compilation of reliable statistics on language pairs and document production will be an important feature in any decision to change current practice.

The final problem in which technological solutions might be appropriate is the need to track stages in workflow efficiently. In such a large and diverse organisation, prioritising and tracking translations is essential at unit level so that Heads of Unit can prioritise tasks, ensure deadlines are met and assign work if necessary. In the past, units have relied on the *Fiche de travail* and paper-based systems such as coloured index cards or

work trays to monitor progress of translation jobs and most units currently maintain this sort of system, in conjunction with electronic tracking via Dossier Manager in the Commission. As the workforce grows and more tasks are completed externally, tracking workflow electronically is likely to become more important across institutions.

Once again, this raises a fundamental dilemma: when do the benefits of increased standardisation in the use of technology justify limitations on diversity? Needless to say, this is key problem for all working in the industry today, and is likely to carry on being a relevant question in the foreseeable future. What other challenges and problems might we also expect, based on the issues identified in the EU institutions?

3. Future challenges

What did interviewees identify as the main challenges in workflow and multilingual document management in the near future?

On the immediate horizon, the introduction of Euramis concordance features and IATE to build on and share terminology resources were generally mentioned as positive developments, though the challenge is apparent in that users will need to adapt to a new interface. Clear instruction leaflets (*Starter Kit: Euramis Concordance*) and online help have been produced and advertised to address initial unfamiliarity. The design of the interface has also been closely modelled on Eurodicautom to ease users in gently. This sort of inter-institutional collaboration is likely to become more common in future, and with such collaboration will come increased challenges in communicating and adopting best practice across larger institutions. Effective communication using technology is likely to be key here, with the institutions sharing more information and communicating via intranet, for example.

Also related to enlargement is the move to outsourcing more translation work. This raises further potential challenges, notably how far tools and information can be shared to maintain quality while safeguarding confidentiality. Might freelance translators return TMs for checking and uploading to the central TM? Building up electronic resources in the new official languages of the EU is a current challenge which might be partially addressed in this way.

The current tender process for a TM tool is raising new challenges. Simply evaluating the different tools is proving time-consuming at a hectic point in the institutions' work cycle. Once a tool is selected, the challenge of integrating it in existing EU workflow models without dramatic changes to translators' working methods will certainly require significant input from the translation tools teams. It will be important to ensure data is safeguarded and available via the new tool. TMX compliance should help here, as all data in Euramis is already saved in this format.

Finally, given the institutional context, it is hardly surprising that many interviewees felt the main challenge in the future was predicting demand for their services. Political imperatives and events mean it is hard to predict how translation needs will evolve, or plan services as a result, particularly now ten new members have just joined the Union. Their influence on working methods and languages used in the institutions may take some time to be felt but is sure to have an impact. Just as earlier waves of accession countries brought new skills and ways of working to translation services, so the new members are likely to propose innovative solutions and fresh challenges, making the EU a fascinating and rewarding focus for study for the foreseeable future.

Appendix 1: Questionnaire design

Research for this paper included guided interviews with translators based on questions in the following areas:

- Typical translation workflow
- Changes to translation workflow
- Use of computer tools
- Efficiency gains
- Division of tasks
- Effects on quality of computer tools
- Common problems
- Future challenges

In addition, tools specialists were asked to address the following topics:

- Training
- Data management
- Efficient use of tools
- Tools currently available

- Likely changes to tools

The author gratefully acknowledges research funding from the University of Leeds and the participation and support of European Union staff, particularly in the Directorate General for Translation of the European Commission.

Bibliography:

Austermühl, Frank (2001) *Electronic tools for translators: Translation practices explained* Vol. 2. Manchester: St Jerome

Bowker, Lynne (2002) *Computer-aided translation technology: A practical introduction* Ottawa: University of Ottawa Press

Directorate General for Translation of the European Commission (2002, 2004 eds)
Translating for a Multilingual Community
www.europa.eu.int/comm/dgs/translation/bookshelf/brochure_en.pdf

Directorate General for Translation of the European Commission (2002, 2004 eds)
Translation Tools and Workflow
www.europa.eu.int/comm/dgs/translation/bookshelf/tools_and_workflow_en.pdf

Esselink (2000) *A practical guide to software localization* Amsterdam and Philadelphia: John Benjamins

Evans, Andrew & Dimitri Theologitis 'Putting the VIM into AVIMA' in Catriona Picken (ed) *Translating and the computer 12, Applying technology to the translation process* London: Aslib, 129-138

Lambert, José (1996) 'Language and translation as management problems: A new task for education' in Cay Dollerup and Vibeke Appel (eds) *Teaching translation and interpreting 3* Amsterdam and Philadelphia: John Benjamins, 271-293

McCallum, Bob (2004) 'Translation technology at the United Nations' in *MultiLingual computing and technology*, Vol. 15, Issue 2, 19-22

Picken, Catriona (1991) 'Introduction' in Catriona Picken (ed) *Translating and the computer 12, Applying technology to the translation process* London: Aslib, vii-viii

Somers, Harold (2003) *Computers and translation: a translator's guide* Amsterdam and Philadelphia: John Benjamins

Taes, Andre (2001) *Introducing CAT in the European Commission's Translation Service*
www.europa.eu.int/comm/translation/reading/articles/pdf/2001_03_30_brussels_taes.pdf

Wagner, Emma (ND) 'Contacts between universities and the EU translation services – Successes and failures' in Belinda Maia, Johann Haller and Margherita Ulrych (eds), *Training the language services provider for the new millennium – Proceedings of the III encontros de tradução de AsTra-FLUP* Porto: Faculdade de Letras, 397-405

Wagner, Emma, Svend Bech & Jesús Martínez (2001) *Translating for the European Union institutions* Manchester: St Jerome

Williams, Jenny & Andrew Chesterman (2002) *The Map: A beginner's guide to doing research in translation studies* Manchester: St Jerome

Wilson, Barry (1991) 'GEPRO – a computerised system for controlling the production of translations in the European Parliament' in Catriona Picken (ed) *Translating and the computer 12, Applying technology to the translation process* London: Aslib, 143-148

www.europa.eu.int

ⁱ Most of the institutions have made TWB available to all in-house translators since 1997, but there are some exceptions such as the European Parliament. Where this article refers to a particular institution such as the European Commission, this is to indicate that not all institutions work in the same way on the given point.

ⁱⁱ AVIMA was an acronym for Avis de marché, the French term for a notice of invitation to tender, the document type for which AVIMA aimed to rationalise the translation process. See Evans & Theologitis (1991) for further details. GEPRO was an early computerised system for controlling translation workflow in the European Parliament. See Wilson (1991) for further details.

ⁱⁱⁱ Translations outsourced to freelance agencies are not included here for reasons of space; the pattern of workflow is however similar, with pre-processing done in-house and an additional later stage of quality checks and feedback, generally carried out in the relevant language units.

^{iv} EU translation services use the term 'document' to refer to texts for translation, whether this implies a short paper-based source text or multiple related electronic files in varying formats. In practice, the majority of documents currently arrive as MS Word files.

^v Other units, for example the Parliament, use the term *Feuille de route* to refer to a similar cover sheet.

^{vi} Euramis is an acronym for European Advanced Multilingual Information System. Its goals and key features are outlined later in this section.

^{vii} By 2004, approximately 260 translators in the European Commission were using Dragon software in the six official languages in which it is currently available. Source: *Translation Tools and Workflow*: 13.

^{viii} The following source languages are supported: English, French, German, Greek, Spanish. In addition, some other target languages are available and there are some prototype versions available for further language pairs. Source: *Translation Tools and Workflow*: 12.

^{ix} Source: *Commission adopts measures to match supply and demand for translation*. Press release, Brussels, 26 May 2004, available online at www.europa.eu.int.

^x Secretarial staff must pass Entry Competitions in three official languages.

^{xi} Source: *Commission adopts measures to match supply and demand for translation*. Press release, Brussels, 26 May 2004, available online at www.europa.eu.int.

^{xii} See *Council Regulation No1 of 1958 determining the languages to be used by the European Economic Community*: Article 2: 'Documents which a Member State or a person subject to the jurisdiction of a Member State sends to institutions of the Community may be drafted in any one of the official languages selected by the sender. The reply shall be drafted in the same language'; and *Treaty establishing the European Community*: Article 21: 'Every citizen of the Union may write to any of the institutions or bodies referred to in this article or article 7 in one of the languages mentioned in Article 314 and have an answer in the same language' [that is, any of the EU's official languages].

^{xiii} Source: interview with translation tools unit staff, DG for Translation, European Commission.

^{xiv} Source: presentation by Rebecca West (then Manager of external translations at the European Parliament), ITI Edinburgh conference on Translation Quality, March 2004.

^{xv} EU in-house translators can apply to work from home for fixed periods and access helpdesk support by phone or email.

^{xvi} See McCallum (2004: 19-22) for a discussion of similar problems posed by organisation size with reference to translation technology at the UN.