

Computer tools to improve the work of translators

Nigel Cassar

Documentation, Reference and Terminology Section, Translation Division, Department of Conference Services, United Nations, New York

In the course of work, staff translators in large organisations are beset by a variety of difficulties: lack of suitable equipment, inadequate sources of information, production difficulties, impossible deadlines, cumbersome institutional vocabulary and jargon, to name but a few. At the United Nations, the entire documentation and translation delivery process is being reviewed and a new system is in the course of development. A number of sub-modules are in the course of installation and further subsystems will be brought in during the next two budget two-year cycles. It is proposed to complete development work on the seven language terminology information system, making it accessible to all translators throughout the entire UN system of organisations. A direct connection between the online terminology system and the bibliographical information system has been established and a further connection is being developed to a full text optical disk system, currently being tested on a worldwide basis. The system uses existing communications technology to link up offices scattered around the globe. The expectation is that the translators will all have direct access through translator workstations to text and data systems and be able to perform on-screen translation. Other areas are intelligent subsystems used in language recognition and text input, multilingual word processing, and multilingual text databases. In the long-term perspective we are interested in voice-activated systems which could be hooked up to machine translation systems. Machine and machine-assisted translation is in use in two of the agencies of the UN system and virtually all agencies have performed tests of various sorts at this stage. It is anticipated that there will be operating systems in most agencies to cover the highly repetitive and routine technical documents, while leaving the complicated and difficult materials for machine-assisted

means, including full revision and final copy review. An implementation programme is under way and within the next six to eight years many aspects of documentation processing and information sharing will have been automated in such a way that the translation process will incorporate many modules that will solve most of the difficulties encountered by today's translators.

At the United Nations, in New York, we perform translation services in two main areas. Firstly, in our main stream of documentation, we process each year some 85,000-95,000 pages of parliamentary records, institutional documents, proceedings of technical bodies, reports and committee papers. These documents are delivered for translation in a number of different languages but are primarily submitted in either English or French and to a certain degree in Spanish, with some of the work from the regional offices and national governments delivered in other languages. The official languages of the major bodies are, in the order of adoption by the United Nations, English and French, then Spanish, Russian, Chinese and Arabic. There is some translation into German of General Assembly documents for the four German speaking member states. The official policy in the processing of the languages is that they all have equal treatment and are processed on the basis of complete parity. There are certain complications in the actual processing of the different languages either because of aspects that are inherent to a particular language or because of the technical means used to process any particular language. In some cases there are factors which tend to slow down or speed up production at different stages such as linguistic reasons – staff availability, operational constraints or technical aspects. Secondly, in the other area of documentation, there are approximately 40,000 pages of special technical papers, organisation records, substantive documentation, project reports, statistical materials, publicity materials and documents such as texts of international treaties which are registered with the United Nations. Once again, the original documents are usually but not always submitted in the official languages with a vast majority of internally generated documentation being written in either English or French. The majority in New York is written in English, roughly 80 per cent of the workload. The international treaties are submitted in a great variety of languages. With the exception of the summary and verbatim records which are written by language staff, a lot of the writing is done by non-native speakers and, as a result, the majority of documents are edited generally for style, grammar and spelling but not for content. Translations are required for both types of documents and in general all of the first group are translated internally by staff translators. The vast majority of the documents in the second group are translated externally by freelancers who are usually retired staff translators. The gross number of pages to be processed in any one language would therefore be approximately 130,000-135,000 units and average costs across all languages per word would average out at approximately 12-17 cents depending on language combination or direction and degree of textual complexity, with staff costs running at about \$40 per hour without overhead costs and a circulation of some 5,000 copies of

each page. In terms of sheer processing costs and performance, the figures sound massive but the products are delivered on time every day and documents can be processed almost any time of day or night more or less like a commercial newspaper production line. Virtually all the document processing line is broken up into shifts and we have coverage pretty well all around the clock in certain areas such as printing and through the weekends as well; consequently we can cover any emergency or disaster situation every single day of the year. A major disadvantage in terms of costs is the overhead factor for a site such as New York, and in particular, for any operations conducted off-headquarters premises. However, the advantage of being able to congregate representatives of all the member states in one location outweighs this disadvantage and the site is strategically convenient in that it is not in a capital city but nonetheless in a major communications and travel zone.

In the Translation Division we are undertaking an internal technological innovations programme, which has started in the Documentation, Reference and Terminology Section. The section receives most manuscripts before they go to the translators. We check the manuscripts for parts of texts which have been previously translated. In general, these parts of texts are quotations whether identified or not, titles of agreements and institutional nomenclature. Any other background documents which already exist in the target languages are also passed on to the translators as background. In some cases, pages and pages are repeated from parent documents or are quotations of legal instruments which have to be retaken word-for-word from the originals. The savings in translator and reviser time are enormous if the previous translation can be used. Another consideration is that we are obliged to use the existing authorised or accepted documentary sources for all of the languages when working on the texts of new legal instruments. In the long term, we would like to receive manuscripts only in electronic form as word processing documents since this will help us to match portions of texts. At the other end of the production line we have our terminologists who scan all documents in hard copy in all the language equivalents as they are published and they record all the language usage of the organisation in the various sectors of activity. Once again, we would like to receive the final target language documents only in electronic form as word processing documents and this will improve terminology data collection.

With regard to technology, we have focused attention so far on improvements that we could make at division level which would link into other areas which are also being automated. Fortunately, all of the areas we have concentrated on are relatively easy in crude technical terms. My own section specifically covers documentation, reference and terminology and it also includes a specialised library for the translators. In each area of activity we have implemented a separate system to automate part, if not all, of the functions undertaken. So far, the results have been good because we can now do more sophisticated research, we can be more imaginative in our data retrieval strategies and we can find large amounts of information very fast; we can also

find out when we don't have any information, which is extremely important, and we can know with certainty and very much quicker if information is not actually available. We are now able to process far more documentation because of our information sources and, consequently, our production figures have risen in our own area from about 145,000 pages per year to about 230,000 pages aggregate per year for all languages with a loss of about three staff posts.

In my area, we have introduced word processing and data processing equipment on an incremental basis and most of the staff now compete to use or share equipment, making it clear that full automation should proceed so that all staff have their own equipment where it is required. Not everything has gone as we would have liked: we had hoped to have an online multilingual terminology databank in all the official languages some time ago but there have been massive technical problems. These arose because of the combination of Roman and non-Roman scripts used in the particular language combination that we use at the UN. From the end of 1972, we have had partial solutions and a number of subsystems or mixtures of systems, using real characters for one language, transcriptions for another or a mainframe dataprocessor for three or four languages and then typing or word processing for the others. The mere fact of mixing automated systems with traditional systems became a nightmare, since material in some languages could always be retrieved and updated in minutes and in large quantities, while languages like Chinese and Arabic were being processed by traditional means with ever-increasing delays as the volume of machine-stored data increased. By 1983, we had gone from producing small terminology bulletins with perhaps 300 terms in English, French and Spanish, with indexes to the respective languages, to a system where we could easily produce listings of 3,000 terms in the original languages together with Russian by this stage, but the Chinese and Arabic were being collected, compiled, numbered, typed, indexed and retyped, all by traditional means. The extra planning was impossible after a certain level since the delays were too massive: it was clear that a full multilingual system was needed.

We were very lucky some four years ago, when a Japanese programmer came to us on loan from IBM in Japan and through his dedication we now have a unique system: it consists of equipment loaned by IBM Japan and New York, with pieces of laser print systems scrounged from Beijing, Brussels and Geneva; we use normal American and Canadian terminals with multilingual keyboards and a Japanese IBM PC; some of the software came from Tokyo, the actual programs have all been re-written internally. We now have a system that will show all of the official languages in one display on one type of screen by using a special double eight-bit system. We can also search the different languages interactively either directly or through transcription depending on what kind of terminal is available and the system is powerful enough to take combinations of data from a number of data fields. The quality of the printed output is also excellent since it is laser printed from a mainframe-driven system and it is camera-ready if we wish to produce terminology publications (see Annex).

This mixture of equipment has been the major breakthrough for us in terminology and it will improve our storage and production capabilities. It will also give the possibility of establishing a proper databank and we believe that we can make this system available to all our colleagues within the UN system of organisations and our counterparts in the major international organisations. Many of the agencies have established terminology databanks and agreements exist for full collaboration at inter-agency level. I should add here that we have been using a number of the national terminology databanks with great success and some of our agencies have been hooked up to the European Eurodicautom system for many years. In one case, the Canadian Termium system, we now have access on two CD-ROM players and this has changed things quite dramatically since we are now able to have access at any time both in New York and Geneva or dial in direct to Ottawa through telecommunications or a telephone tie line. We will clearly be going in that direction ourselves since it seems to be the most efficient way of sharing out data to remote locations. For a number of years, we experimented with copying data on tape and even more recently on word processing disks and sent out the data to various agencies but this proved clumsy and required considerable maintenance work.

Having achieved full 'multilinguality', we can now turn to the provision of full systems for other applications where so far we have been coping with only one or two languages. In other areas, for instance, we would like to change the language coverage in our division library where the bibliographical and cataloguing system should be linked through a common multilingual thesaurus to the terminology database. At the same time, we would like to bring it into line with the main library which has a major bibliographical system. That system is called UNBIS and is shared by a number of our offices and could serve as a central indexing engine to locate all documents and activities from anywhere within the United Nations. At present, it cannot represent all the official languages in the database, which is why it is critical for us to merge these two areas. In the area of documentation, we are looking at a massive storage system to replace our old microfilm and microfiche systems. The UN Geneva office is using our section as part of a test of a major documentation storage and delivery system in collaboration with the French SARDE optical disk system. Since that system requires an automated index to drive it, we anticipate that the same multilingual thesaurus and UNBIS index will bring us to the required documentation. Another information system of direct use to us is the UN Treaty Information System and it is hoped that it can be brought into the overall picture. Most of these systems are bibliographical research and documentation retrieval tools. The translators, the terminologists and the research clerks will be able to use all of these.

At another level, we are automating the maintenance of production records and the tracking of documents as they pass through the entire production chain. We are hopeful that the system will be fully operational by the end of the year. We expect that additional refinements will be added during the coming year. It

might be possible to have documents automatically classified and indexed either through bar code labels or at source as soon as documents enter the first level of automation when they are typed in electronic form. The present tracking system will follow a publication all the way through to its final issuance and even indicate international standard serial book numbers, price and other information. Similar database systems are being created to program and keep track of conferences, interpreter availability and assign resources from one meeting area to another.


As an inter-agency project, John Alvey at the World Bank is developing an automated directory of translation and terminology resources which can be accessed directly and I am creating a subsystem to include a list of in-house specialists in the UN and contacts who could be of use to the translators. I would like to see this extended eventually to include other resources. I am also interested in developing contacts at the national level with government agencies, reference centres and in particular with academic and language studies bodies. I should add in passing that we have been running a successful student internship programme for the last few years.


In conclusion, we would like to provide translators with workstations on an elective basis, and they will have direct links to terminology, reference, legislative and bibliographical sources; we would like the translators to word process documents under five pages long and these will be transmitted to the typing pools for final copy preparation, otherwise longer documents will continue to be dictated and transcribed by professional typists. Our major problem at present is that although we have the different pieces of an overall system in place, they are not all available from the one device, and our elusive translator workstation still does not have access to everything we require. We are therefore concentrating during the course of the next few months on working with a couple of major equipment and system vendors and we hope to get some results fairly soon. Our ultimate goal is still to see some measure of automated translation but results in our particular area have not been positive: a special area where there could be an application is for the notification of space launches which are routine information notes but which barely constitute 2,000 pages of documentation each year and do not justify operation of an entire system.


AUTHOR


Nigel Cassar, Chief, Documentation, Reference and Terminology Section, Translation Division, Room 1153B, Department of Conference Services, United Nations, New York NY 10017, USA.

Annex: Sample Pages

 TERMINOLOGY LIST	
1	<p>Abidjan Statement Déclaration d'Abidjan Declaración de Abidjan Абиджанское заявление [E/ICEF/1987/2; para 41 (24 February 87)] [UNDP; All-Africa Seminar on Low-Cost Rural and Urban-Fringe Water Supply, Abidjan, Oct. 1986] [24 February 88 rb/dlg]</p>
2	<p>ABM Treaty 反導彈條約 [Disarmament Bull., No. 606] [see: Treaty between the United States of America and the Union of Soviet Socialist Republics on the Limitation of Anti-Ballistic Missile Systems] [11 Feb 87 HV]</p>
3	<p>Abstracts of Selected Solar Energy Technology; ASSET/UNU Analyse de certaines techniques relatives à l'énergie solaire; ASSET/UNU Resúmenes de determinadas tecnologías de energía solar; ASSET/UNU 太阳能技术选萃 علامات بحوث مختارة في تكنولوجيات الطاقة الشمسية</p>
4	<p>Abu Dhabi Fund for Arab Economic Development Fonds d'Abou Dhabi pour le développement économique arabe Fondo de Abu Dhabi para el Desarrollo Económico Árabe 阿拉伯经济开发基金 مندوق ابوظهبي لتنمية الاقتصاد العربي</p>
5	<p>Abuja Statement on Economic Recovery and Long-term Development in Africa Déclaration d'Abuja sur le redressement économique et le développement à long terme de l'Afrique Declaración de Abuja sobre la recuperación económica y el desarrollo a largo plazo de África Абужийское заявление об экономическом подъеме и долгосрочном развитии в Африке A/42/410 Annex (21 July 87) 关于非洲经济复苏和长期发展的阿布贾声明 بيان ابوجا بشأن الانتعاش الاقتصادي والتنمية الطويلة الاجل في افريقيا [A/42/410 Annex (21 July 87)] [issued at International Conference on Africa, Abuja, Nigeria, 15-19 June 87] [4 August 87 REB]</p>
6	<p>Academic Policy Committee Comité chargé de définir les grandes orientations de l'Université Comité de Política Académica 学术政策委员会 لجنة السياسة الأكاديمية [UNU]</p>
7	<p>Academy of Criminal Justice Sciences; ACJS Académie des sciences de criminologie Academia de la Ciencia de la Justicia Penal Академия криминологии; АК [E/C.2/1989/2, p. 5 of E; YIO GG4032] [US-based NGO] [13 Feb 89 rb/dlg]</p>
8	<p>Acapulco Commitment to Peace, Development and Democracy Engagement d'Acapulco en faveur de la paix, du développement et de la démocratie Compromiso de Acapulco para la Paz, el Desarrollo y la Democracia Акапульское обязательство в отношении мира, развития и демократии [A/42/844 (1 December 87)] [Signed by Argentina, Brazil, Columbia, Mexico, Panama, Peru, Uruguay, and Venezuela, 29 November 87] [10 December 87 rb/dlg]</p>

 TERMINOLOGY LIST	
9	<p>Acapulco Document on World Tourism Document d'Acapulco sur le tourisme mondial Documento de Acapulco sobre el Turismo Mundial 关于世界旅游问题的阿卡普尔科文件 وثيقة أكابولكو بشأن السياحة العالمية [Adopted by the World Tourism Meeting; Acapulco, México; 21-27 August 1982]</p>
10	<p>ACC Inter-Agency Task Force on Information Exchange and the Transfer of Technology: IATFIS Equipe spéciale interorganisations pour l'échange d'informations et le transfert des techniques Equipo de tarea entre organismos sobre intercambio de información y transmisión de tecnología 行政协调会资料交换和技术转让机构的工作队 فرقة العمل المشتركة بين الوكالات التابعة للجنة التنسيق الإدارية والمعنية بتبادل المعلومات ونقل التكنولوجيا [ACC]</p>
11	<p>ACC Intersecretariat Group for Water Resources Groupe intersecrétariats pour les ressources en eau (mis en place par le CAC) Grupo Intersecretarial sobre Recursos Hídricos Межсекретариатская группа АКК по водным ресурсам [E/1987/21 (Supplement 8 of ECOSOC), para 99; ACC/1987/1, para.35(b)] [19 August 87 REB]</p>
12	<p>ACC Sub-Committee on Water Resources Development Sous-Comité de la mise en valeur des ressources hydrauliques Subcomité sobre el Desarrollo de los Recursos Hídricos 行政协调会水力资源发展小组委员会 اللجنة الفرعية لتنمية الموارد المائية التابعة للجنة التنسيق الإدارية [ACC]</p>
13	<p>ACC Task Force on Long-term Development Objectives Equipe spéciale [du CAC] sur les objectifs de développement à long terme Equipo de Tarea [del CAC] sobre Objetivos a Largo Plazo Целевая группа АКК по долгосрочным целям развития [E/AC.51/1988/7; Table 3; ACC/1986/87; F; groupe spécial sur les ...] [17 May 88 rb/dlg]</p>
14	<p>ACC Task Force on Science and Technology for Development Equipe spéciale sur la science et la technique au service du développement Grupo de Trabajo sobre Ciencia y Tecnología para el Desarrollo 行政协调会科学和技术促进发展工作队 فرقة العمل المعنية بتفسير العلم والتكنولوجيا لأغراض التنمية التابعة للجنة التنسيق الإدارية [ACC]</p>
15	<p>Access to Sources of Knowledge: ASK Acceso a Fuentes de Información 接触知识来源的机会 الوصول إلى مصادر المعرفة Access to Sources of Knowledge: ASK; consultation des sources [information referral service title] d'information</p>
16	<p>Accident Prevention Programme Programme de prévention des accidents Programa de Prevención de Accidentes Программа предупреждения несчастных случаев 预防事故方案 برنامج الوقاية من الحوادث [WHO]</p>

 TERMINOLOGY INDEX LIST	
Административная служба Administrative Service Service administratif	403
Административная служба связи со средствами массовой информации Executive Media Service Service des médias/Haute direction	4391
административное задержание administrative detention Internement administratif	399
Административно-финансовый отдел Division of Administration and Finance Division de l'administration et des finances	3721
Административно-управленческий отдел Administrative and Management Services Division des services administratifs et de gestion	391
Административный комитет по вопросам борьбы со СПИД AIDS Management Committee Comité de gestion pour le SIDA	841
Административный отдел Division of Administration Division de l'administration	3719
Административный совет Administrative Council Conseil d'administration	398
Афинская резолюция Athens resolution résolution d'Athènes	1435
Африканская биологическая сеть African Biological Network réseau biologique africain	528
Африканская комиссия по правам человека и народов African Commission on Human and Peoples' Rights Commission africaine des droits de l'homme et des peuples	543
Африканская компания по управленческому обслуживанию African Management Services Company; AMSC Société de services de gestion pour l'Afrique; AMSC	614
Африканская конференция парламентариев по вопросам народонаселения и развития African Conference of Parliamentarians on Population and Development Conférence africaine des parlementaires sur la population et le développement	549
Африканская конференция по диарейным заболеваниям African Conference on Diarrhoeal Diseases Conférence africaine sur les maladies diarrhéiques	552

 TERMINOLOGY INDEX LIST	
<p>الفريق الاستشاري المخصص لموضوع التفجيرات النووية للأغراض السلمية Ad Hoc Advisory Group on Nuclear Explosions for Peaceful Purposes Groupe consultatif ad hoc sur les explosions nucléaires à des fins pacifiques Grupo Asesor Especial sobre Explosiones Nucleares con Fines Pacificas</p>	55
<p>الفريق الاستشاري المخصص للشباب Ad Hoc Advisory Group on Youth Groupe consultatif spécial pour la jeunesse Grupo Asesor Especial sobre la Juventud</p>	56
<p>اللجنة الاستشارية المخصصة لبرامج البحوث والتدريب في التنمية الإقليمية Ad Hoc Advisory Committee for the Research and Training Programme in Regional Development Comité consultatif spécial pour le programme de recherche et de formation portant sur le développement régional Comité Asesor Especial para el programa de investigación y capacitación en la esfera del desarrollo regional</p>	54
<p>اللجنة الفرعية لتنمية الموارد المائية التابعة للجنة التنسيق الإدارية ACC Sub-Committee on Water Resources Development Sous-Comité de la mise en valeur des ressources hydrauliques Subcomité sobre el Desarrollo de los Recursos Hídricos</p>	12
<p>اللجنة المخصصة المنشأة بموجب قرار مجلس الأمن Ad Hoc Committee established under Security Council Resolution 507 (1982) Comité spécial créé par en application de la résolution 507 (1982) du Conseil de sécurité Comité ad hoc establecido por la resolución 507 (1982) del Consejo de Seguridad</p>	60
<p>اللجنة المخصصة المنشأة بموجب قرار ٤٥٥ (١٩٧٩) والخاصة بشكوى زامبيا، لمساعدة مجلس الأمن في تنفيذ القرار Ad Hoc Committee Established Under Resolution 455 (1979), Concerning the Complaint by Zambia, to Assist the Security Council in the Implementation of the Resolution Comité spécial du Conseil de sécurité créé par la résolution 455 (1979), relative à la plainte de la Zambie, pour aider le Conseil de sécurité à appliquer ladite résolution Comité ad hoc del Consejo de Seguridad establecido en virtud de la resolución 455 (1979), relativa a la denuncia de Zambia, para asistir al Consejo en la aplicación de la resolución</p>	59
<p>الممثل المقيم بالنيابة Acting Resident Representative Représentant résident par intérim Representante Residente interino</p>	35
<p>الرئيس بالنيابة Acting Presiding Officer Président par intérim Presidente interino, Presidenta interina</p>	34
<p>الوصول الى مصادر المعرفة Access to Sources of Knowledge: ASK Acceso a Fuentes de Información</p>	15
<p>برنامج العمل الخاص بالمخاقير الأساسية Action Programme on Essential Drugs Programme d'action pour les médicaments essentiels Programa de acción sobre medicamentos esenciales</p>	49