

Terminology management in NATO

Ian JONES, SHAPE, Belgium
jonesian@shape.nato.int

Introduction

The North Atlantic Treaty Organization (NATO) now has nineteen member states which take part in its political decision-making bodies. Most of the members also contribute military personnel to the NATO command structure and military forces to NATO's permanent structures or to ad hoc coalitions such those currently operating in Bosnia-Herzegovina and Kosovo. The main elements of the organization itself are:

- ◆ the International Staff, which acts as the Secretariat of the North Atlantic Council, the supreme political body governing NATO and its multiple subordinate committees,
- ◆ an International Military Staff, which acts as the Secretariat of the NATO Military Committee which reports to the North Atlantic Council and is composed of the Chiefs of Staff of the various member nations,
- ◆ two strategic commands, one in the United States and one in Europe, which in turn have a number of subordinate commands in various parts of the NATO area,
- ◆ a number of specialized agencies operating in various fields: research and technology, support for weapon systems, logistic support, standardization, etc.

Furthermore, NATO is associated with 25 other countries in the "Partnership for Peace" programme, making a total of 44 nations which are endeavouring to cooperate not only on a political level but also in exercises, planning and real operations such as those in the Balkans. As well as "PfP" nations, other countries, such as Argentina, Malaysia, Morocco, etc. have contributed or are contributing to operations in the former Yugoslavia.

The languages of NATO

NATO's official languages are English and French. However, the various bodies and headquarters located throughout the NATO area and outside it also have a requirement for their host country languages. Lastly, the Alliance is making great efforts to maintain friendly relations with countries in central and eastern Europe, leading to a requirement for Russian amongst other languages.

Terminology in NATO

The reason I mention the make-up of NATO and the languages used is to illustrate the diversity of our requirements for terminology. Terms need to be managed for purposes we are familiar with: translation and interpretation, in which case the requirement is for accurate and consistent translation. But control of terminology is also important for equipment and document production, for standardization and for what we call the "operators", i.e. military personnel planning and conducting operations. When coalitions of forces are put together from the various services and nations, there has to be a common language used and understood by all to prevent confusion and misunderstandings which can lead to the loss of lives or even battles. Modern equipment programmes are highly complex, involving contractors and sub-contractors in a variety of countries who must be able to understand each other unambiguously even when they are all using the same language.

Changing world - changing terminology

As well as the diversity of our requirements, we are faced with the need to take account of accelerating changes in the political and strategic environment in which we are now living. During the years of the Cold War, that environment changed relatively little and when it did change, it did so quite slowly, so that traditional methods, such as publishing glossaries in paper form, were adequate. Now we suddenly find ourselves needing to define terms like *enabling forces*, which prepare the ground for *peace making, peace building, peace enforcement* or *peacekeeping* forces. We talk about *information warfare, psychological operations, civil-military cooperation* and *humanitarian operations*. NATO needs to be able to distinguish between *refugees, displaced persons* and *evacuees*. We must prepare for military operations conducted with non-NATO members in *Combined Joint Task Forces*. The NATO military command structure has just been reorganized with the result that we no longer have *Major NATO Commands, Major Subordinate Commands* and *Principal Subordinate Commands*, but *Strategic Commands, Regional Commands* and *Joint Sub-Regional Commands* or *Component Commands*...

Military Agency for Standardization

NATO has an organization called the Military Agency for Standardization (MAS) whose job is lay down norms for equipment, communications and the military doctrines applied for the employment of various kinds of forces: land, maritime, amphibious, air, joint, etc. For this purpose it has a network of committees of national and NATO experts dealing with standardization in particular subject areas. As part of their work, they produce documents which often incorporate specialized glossaries on many subjects ranging from map-making to medical services, from logistics to electronic warfare. These documents and terminology are adopted officially by the member countries and are thus normative. MAS also has a committee called the NATO Terminology Conference which is responsible for coordinating terminology in general and more particularly produces two Allied Administrative Publications: AAP-6, the *NATO Glossary of Terms and Definitions* and AAP-15, the *Glossary of Abbreviations used In NATO documents and*

publications. AAP-6 is updated continuously and AAP-15 is revised at regular intervals. These documents can now be consulted on the NATO web site (www.nato.int).

NATO Linguistic Services

The NATO linguistic services also record and exchange terminology for the purposes of harmonizing translation and interpretation services. This terminology is controlled by the linguists themselves and is considered as a working tool for linguists. It is not therefore normative. However it has a considerable impact on both the quality and efficiency of translation services in particular.

The terminology referred to above is produced in NATO's official languages: English and French. However, much of it is also translated into other languages either by linguists working in NATO bodies where the host country language is neither English nor French, or by linguists working for ministries in those countries.

Other sources

NATO also makes maximum use of terminology adopted and promulgated by other international organizations: ISO, IEC, ICAO, etc.

Choice of system

The reader will grasp from the above that all this terminology exists in a variety of forms and formats, comes from a variety of sources and is required for a variety of different purposes. We consider that those purposes, while not identical, overlap sufficiently to justify trying to find a method of bringing together all that information and making it available in a form that it is usable and useful to the widest possible community.

Translators require a system integrated into their other software tools and which provides them with accurate translations. Equipment producers need to be able to ensure that all their suppliers and customers understand specifications, manuals, etc. Military officers writing operational plans need to be sure that units implementing plans understand their missions. Air traffic controllers want to be sure that pilots have understood instructions, and so on. Lastly, the system has to be capable of incorporating and disseminating changes in terminology rapidly and efficiently.

Furthermore, all this has to be done with limited resources: our budgets and manpower are tightly controlled. We therefore have to find an affordable solution that is flexible enough to meet differing requirements. We must also find a suitable off-the-shelf system as we have neither the time, the expertise, nor the money to develop custom-made systems.

In this area, the lead has been taken by a number of NATO linguistic services which began several years ago to adopt translation memory systems such as the IBM Translation Manager or the TRADOS Translator's Workbench. To begin with

there was little coordination. Subsequently, the majority of linguistic services decided to adopt the TRADOS system on grounds of its features, flexibility and, at the time, price (it was the cheapest of several similar systems we looked at some years ago).

Multiterm

One of the two major components of the TRADOS Translator's Workbench is the Multiterm terminology data base system. Although, NATO's official languages are English and French, we need a system that allows other languages to be added, including those with non-Latin alphabets, and permits different presentations of the same data to meet the needs of different users. We also require a system that allows us to import data from a variety of sources: Word or WordPerfect files, Access, Clipper, etc. The terminology management system also has to remain integrated with the translation memory system we use and permit the exchange of data among the various organizations.

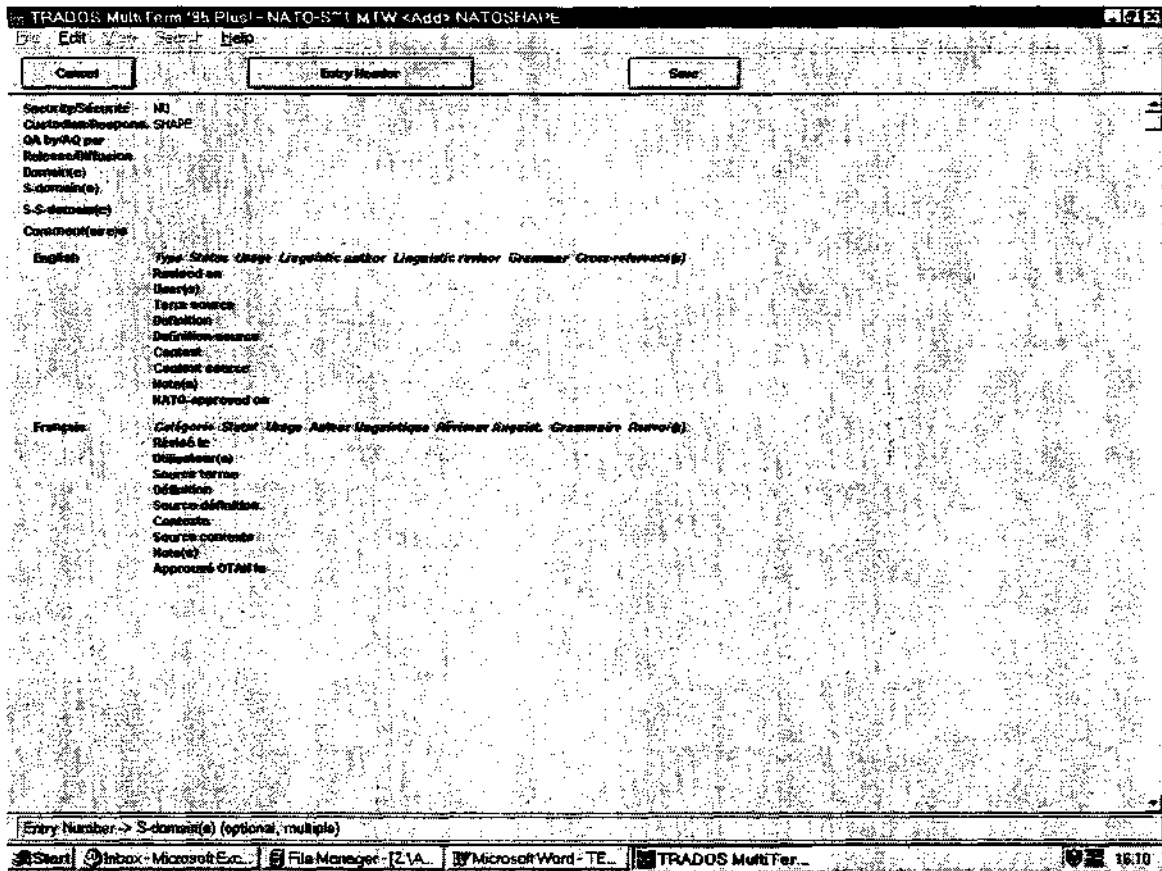
We have found that Multiterm, while not perfect, as I shall explain later, generally meets our requirements.

Standard input models

The Multiterm system allows the user to define the layout of the data it manages: the field names, their order and other attributes. This is known as an input model. To begin with, Multiterm was used by the NATO linguistic services only and we decided to create a standard input model with the fields we felt we required to manage our data effectively. This was not an easy task and required several meetings with sometimes heated exchanges of views! I shall not go into detail on the form that input model took because it has recently been superseded by a new one.

The new model was produced as a result of new participants adopting the system, in particular the NATO Terminology Coordinator and certain member countries, especially Canada which has the same official languages as NATO and huge translation and terminology control requirements. It meant changing the names of certain fields we had already adopted, adding others and defining the contents of each field so as to meet the requirements not only of translators and interpreters but also other actual or potential users of the system.

The input model adopted looks like this:



The fields at the top left hand side of the model relate to the complete record:

- Security/Sécurité
- Custodian/Respons.
- QA by/AQ par
- Release/Diffusion
- Domain(e)
- S-domain(e)
- S-s-domain(e)
- Comment(aire)s

The first four are used by the terminology managers to indicate whether information is classified, the organization which has created the record, the person who has performed quality control on the record and whether it can be released for wider circulation.

The next three are used to indicate the domain, sub-domain and sub-sub-domain to which a record relates. Because of the differing requirements of the users, we have adopted only nine basic standardized domains for all users. These are as follows:

- General Terminology / Terminologie générale
- Political Affairs / Affaires politiques

Defence / Défense
Human Resources / Ressources humaines
Budget and Finance / Budget et finances
Law / Droit
Science
Technology / Technologie
Organizations / Organisations

Users are free to define their own lists of sub- and sub-sub-domains to suit their own purposes.

"Comment(aire)s" is used for any comment concerning the whole record.

For those not familiar with Multiterm, it should be noted that it is a "uninotional" data base, i.e. each record covers a single notion with as many synonyms as are required, but not homographs.

We next see "English". In Multiterm jargon, this is called an "index field" and is used to input a headword or phrase. The following fields, down to "NATO-approved on" all relate to the term entered in that index field. The user can add as many index fields as are necessary to input synonyms, abbreviations, etc.

In the first field "Type", the user can choose one of the following:

<u>Possible value</u>	<u>Example</u>
Term	North Atlantic Council
Abbreviation	NAC
Short Form	the Council
Code	AC/141 (NG/4-SG/42) (i.e. the code for a NATO committee)

In the next field "Status", the user can choose one of the following:

Proposed
Under review
Approved
Outdated

In the Usage field, the user can choose one or more of the following:

Preferred/A préférer
AVOID/A ÉVITER
Informal/Informel
Trademark/Marque déposée
Regionalism/Régionalisme:

"Linguistic author" and "linguistic reviser" are used to indicate the name of the linguists (translator, interpreter, terminologist, etc. who have entered and revised the data).

"Grammar" can be used to indicate the part of speech and other information such as irregular forms.

"Cross-references" is used to refer to terms with related meanings, antonyms, etc., but not to synonyms which must appear in the same record.

"Revised on" is the date when the entry was last officially revised.

"User(s)" indicates the organizations, committees, etc. which use the term.

The remaining fields are probably self-explanatory.

The same fields (translated into French) are found under the "Français" index field. Other languages can easily be added to the same input model simply by creating an appropriate index field followed by the same list of fields translated into that language.

Only the following fields are compulsory, which allows the data base to accept information from a variety of sources which do not necessarily contain data related to each field and also provides flexibility when entering data manually as the user can fill in only those fields he/she considers necessary or for which information is available:

(For the complete record:)

Security/Sécurité

Custodian/Respons.

Domain(e)

(Under each index field:)

Type

Status

Term source

User guide

Lastly, to ensure that the system is implemented in as uniform a manner as possible by the various services using it, we have drawn up a set of detailed instructions (in English and French) for entering data in this input model.

Example

This is an example of a completed record where all fields containing data are displayed. It illustrates the flexibility of the system inasmuch as apart from the compulsory fields, information has only been entered in those fields felt to be necessary by the author (in this case myself).

The screenshot shows the TRADOS MultiTerm '95 Plus interface. The window title is "TRADOS MultiTerm '95 Plus - NATO-S*1.MTW <View>". The menu bar includes "File", "Edit", "View", "Search", and "Help". The "Index" is set to "English" and the "Target" is set to "Français". The term being viewed is "Regionalism".

Creation Date: 05.08.1999 - 11:58:56
Created By: super
Change Date: 06.08.1999 - 09:22:09
Changed By: super
Entry Class: 1
Entry Number: 9
Security/Sécurité: NU
Custodian/Responsable: SHAPE
Release/Diffusion: Yes/Oui
Domain(s): Central Terminology / Terminologie générale

English
Regionalism *Type Term Status Approved Usage Preferred/A préféré Linguistic author JJ Grammar None*
Use(s): Linguistic and terminology services
Term source: NATO data base definition 99.1
Definition: Indicates a term which is only used within part of the area inhabited by native speakers of English.
Definition source: JJ
Note(s): Indicate after "Regionalism/Régionalisme" in which area the term is used, e.g. US, Canada, Scotland, etc.

Français
Régionalisme *Catégorie Terme Statut Approuvé Usage Préféré/A préféré Grammaire Non mentionné*
Utilisateur(s): Services linguistiques et terminologiques
Source terme: NATO data base definition 99.1
Définition: Se dit d'un terme qui n'est employé que dans une partie de la zone géographique habitée par des francophones.
Source définition: JJ
Note(s): Indiquer après "Regionalism/Régionalisme" la zone dans laquelle le terme est employé, par exemple Belgique, Canada, France, etc.

The bottom of the screenshot shows the Windows taskbar with the Start button and several open applications: "Inbox - Microsoft Exc...", "Microsoft Access - M...", "File Manager - Z:\A...", "Microsoft Word - TE...", and "TRADOS MultiTer...". The system clock shows 11:06.

Exchange and distribution of data

Having adopted the common input model described above, the participants in the programme will exchange data base files so that they are available locally to linguists, terminologists and other users.

Ultimately, it will also be possible to make the data available in a suitable form on our Intranets and I hope the World Wide Web.

System limitations

The comments below apply to the current version of Multiterm. They concern the most important defects in the system from our point of view.

- In some countries, there is a legal obligation to make the entire system bilingual. However, the language used for the names of the fields ("captions") cannot be selected by the user. This already causes problems in creating a bilingual input model as there is a 19-character limit on the size of field names. We have done our best to get round this problem in our input model by providing bilingual field names at the record level and English or French names under the corresponding index fields. It will easily be seen that a multilingual data base would be impossible to configure if all field names and values had to be provided in every language.
- There are limits on the amount of data that can be entered in each field and just as importantly, on the size of the complete record (currently 32 kb). This can cause problems when handling large numbers of synonyms and translations in a number of languages possibly together with graphics.
- When placing terms in semantic categories (domain - sub-domain - sub-sub-domain), the system does not permit a genuine tree structure to be generated, i.e. with sub-domains relating to a main domain appearing automatically and likewise sub-sub-domains of sub-domains. In fact one could imagine a bottom-up approach where the user would need to select only the sub-sub-domain and the system would automatically generate the correct sub-domain and domain. Having said that, the sheer intellectual difficulty of codifying semantic categories into a neat three-tier structure is enormous and I freely admit that we have simply ducked the problem by only agreeing on the top-level domains.
- A user can only search for a term in one data base at a time, whether using Multiterm alone or in conjunction with the Translator's Workbench. To get round this problem we have had to merge various data bases so as to provide all the information to the translators. It would be better to be able to search in cascade mode with the user able to define the search order.
- Unlike other data bases, e.g. TERMIUM, it is difficult to carry out full text retrieval, i.e. to search other than in the head words of entries.
- Certain editing features could be improved, for example cut or copy & paste can only be performed on one field at a time, so that an index field complete with its subordinate fields cannot be moved or copied in one operation. When two Multiterm data bases are opened simultaneously in separate windows, it is not possible to open edit mode in each data base and use copy or cut & paste (e.g. by click and drag) from one DB to another, whether of a complete record or any part of it. Many global editing operations require the entire data base to be exported, e.g. to a text file, manipulated and then reimported.
- Lastly, when using the Translator's Workbench, which has an automatic search feature for terms in a designated Multiterm data base, the look-up

function does not look for the longest matching expressions before stopping at single-word hits.

Conclusion

One might conclude from the above that Multiterm is a rather poorly designed system. Nothing could be further from the truth: it is a highly flexible tool specifically geared to the needs of managing terminology. But like any other system, it has its strong and weak points. It is possible that some of the latter will be improved in future versions of the system.

I should like to end by reminding you that, as I referred to last year in my talk on managing translation services to maximize efficiency and quality, the most important element in translation or terminology management is not the software or the equipment, but the people who operate it and in particular their competence and experience. The Multiterm tool we have designed is intended to give the users the information they need in a structured way so that they can make informed decisions in their various fields. But however good this tool is, it is up to the human being at the keyboard to use it intelligently. Just remember: **owning a Stradivarius doesn't mean you are Yehudi Menuhin!**

Postscript

Unbeatable value!

This might be an offer you can't refuse: if you use Multiterm and would like to adopt the same or a similar input model, I would be more than happy to provide you (free of charge) with a copy of the NATO standard model. Just send me an e-mail and I will send you the input model and user guide right back!