

THE TERMINOLOGICAL DATABASE – AN ESSENTIAL TOOL FOR EFFECTIVE SPECIALISED COMMUNICATION

Annamaria KILYENI, Georgeta CIOBANU

Politehnica University of Timișoara

1. Introduction

The present paper approaches some of the key aspects related to the compilation and management of a terminological database (also called termbase in our paper), with special reference to a termbase in the fields of architecture and townplanning with terms in English and Romanian, entitled "Bază de date terminologice din domeniile arhitectură - urbanism – limbile engleză și română " (*Terminological Database in the Fields of Architecture and Town Planning – English and Romanian*), compiled within a CNCSIS research grant. The absence of such a database in Romania requires the development of such a terminographic product based on the management of terminological data according to a conceptual system, a clear classification and organization of terms within logically and ontologically related fields.

1.1. Objectives

From among the objectives we mention:

- the need to create an English and Romanian terminological database in the fields of architecture and town planning, considered top priority fields in implementing European language policies within the exchange of European cultural values;
- the development of inter and multidisciplinary research (terminologists, linguists, architects, specialists in town planning, specialists in database design) aimed at establishing a knowledge base including the basic concepts of the main subfields of architecture and town planning, focusing on contemporary architecture, Romanian traditional architecture, preservation and renovation, urban planning, landscape design, green architecture, sustainable housing;
- the choice of appropriate equivalents as a means of expressing architectural values, creating a bridge between our national cultural identity and European culture;
- the compilation of an economical terminographic product, with re-usable information and standard exchange format.

1.2. Line of work

In order to reach the above mentioned objectives, our terminographic product observes the principles and methods currently used in terminology (e.g. Pavel and Nolet: 2001, Cabre: 1998, Wright and Budin:1997), as presented below.

A first step involved a clear delimiting of the field and subfields of interest.

The application of the onomaseological approach helped in establishing the basic concepts and assigning the corresponding terms in the two languages, English and Romanian.

The documentation sources used include the basic relevant categories used in terminographic information and documentation, both hardcopy and electronic medium: dictionaries, glossaries, vocabularies, thesauri, books, textbooks, manuals, journals, proceedings, catalogues, dissertations, terminological databases and databanks; additionally, oral sources, i.e. specialists in the field provided precious information.

The system of concepts used to manage the terminological data required the systematic ordering of entries.

Due emphasis was given to various communication situations and types of users (architects, designers, specialists in preservation and renovation, students in architecture, academic staff involved in teaching architecture and town planning, architecture companies, translators, interpreters, terminologists, technical writers, documentation specialists and information specialists).

The termbase was designed in co-operation with IT specialists.

2. Termbase Design

The starting point of the present research was an English terminological dictionary for architecture and building engineering (Ciobanu: 1978). Designed as a didactic product first of all, it was the only publication of the sort at the time of publication; the terminology referred to the main traditional subfields of architecture - theory of architecture, town planning, architectural styles, interior decoration, building structures. Later on, a mini-dictionary of English and Romanian architecture terms was included in a textbook written for students in architecture (Teodorescu: 2003; Teodorescu, e-book: 2007).

To meet today's requirements of efficiently accessing knowledge in the various subject fields, the electronic medium represents a must. The following aspects were taken into consideration in the termbase design:

2.1. Technical aspects

The technical features necessary for selecting our computer program depended on the electronic data-processing equipment in most existing Romanian work environments. Virtually, most users run programs on IBM-compatible personal computers under MS Windows. Also, the Office package is currently available. Lack of access to dedicated software used for managing terminological data was the main constraint that imposed the use of Access as the program easily available for the types of users mentioned above (even if this does not represent the ideal solution).

On the other hand, Access provides a format that can be exported and converted into other formats. Even terminologists who make use of Multiterm do start their rough termbase version in Access, and afterwards convert it to the most suitable format. Also, considering the number of entries, Access provides the proper level of performance for retrieval purposes and maintains this level even if the termbase might reach larger proportions.

In deciding on our termbase size we had in mind the system-specific entry structure and the corresponding data categories, the expected number of terminological entries and the data storage architecture. The chosen program is configured to detect doublettes and give a warning, which was very helpful in avoiding the recording of double entries.

In order to simplify access to and retrieval of information we selected the tabular arrangement which allows the visual display of one term per sheet.

2.2. Terminological aspects

From among several possibilities we chose the author-defined entry structure, having in view as main users the classes of experts and semi-experts.

In order to make the termbase a user-friendly tool, we initially examined the data categories present in the ISO 12620 standard (ISO 12620 : 1998) and opted for a minimal set of fields: termen titlu limba engleză, termen titlu limba română, domeniu, informații gramaticale, definiție, context, variante, sinonime, note de utilizare, data înregistrării, autorul

actualizării, data actualizării, autorul actualizării. Such a set would have required relationships between fields and would have complicated the queries. Also, lots of essential information on the terms of the termbase was missing. Therefore, we added more fields: entry term, domain, definition, context, variants, synonyms, usage notes, echivalent limba română, domeniu, variante, sinonime, note de utilizare, data înregistrării, autorul înregistrării, data actualizării, autorul actualizării.

2.3. User Interface Aspects

One of the advantages of using the tabular Access format results in easy to read information displayed on-screen, in a clear layout. The data elements of the whole entry, including source and target language, are displayed on a single screen.

3. Data management for the “Terminological Database in the Fields of Architecture and Town Planning – English and Romanian”

The terminological record is the smallest independent unit in any modern electronic terminographic product. The effectiveness of a terminological information tool mainly depends on the type and quality of information contained in it. This has led to an international consensus among terminologists on the basic data fields pertaining to a concept and comprised in a terminological record (Sager: 1990, Cabré: 1998, Picht: 1985). These data fields can be grouped into broader categories, according to the nature of the information they provide about a given concept in a specialized field:

- linguistic information,
- conceptual information,
- pragmatic information,
- administrative (housekeeping) information.

For practical reasons, we shall present the organization of the information in our termbase in the fields of architecture and town planning according to the four broader types of data categories, rather than describe the kind of information in each individual data field separately.

3.1. Linguistic information fields

In this set of fields the *entry term* (in English), e.g. *arcade*, *green building*, *High Tech*, *historic preservation*, etc. is the key item. Since nowadays terminology compilation is firmly corpus - based, we used a wide range of original documentation sources in the fields of architecture and town planning in order to collect the terms belonging to this particular domain: dictionaries, glossaries, vocabularies, thesauri, books, textbooks, manuals, lecture notes, articles from journals and magazines, proceedings, catalogues, dissertations, curricula, terminological databases and databanks, many of which were consulted online, on the Internet. It is true that one should be cautious about the quality of the materials available on the Internet and always make sure the information they contain is accurate and reliable. Nevertheless, online sources clearly present a number of major benefits over conventional information media: they can be easily accessed, are up-to-date and provide the possibility of selecting the terms and analyzing their usage in context.

Besides all these written documentation sources, we also used oral sources. Some of the information was provided by subject field experts, especially when written sources were scarce. These oral sources were particularly useful in the case of the Romanian terms, listed under another data field, namely *echivalent limba română* (Romanian equivalent).

Despite the theoretical principle of univocity in terminology, which would indeed greatly facilitate communication in specialized fields, things are quite different in practice and most

terminologists now accept the occurrence of synonyms and variants, since one concept may have two or even more lexical representations. Throughout our research we encountered such situations, in which different terms were associated with the same concept. Consequently, all terms matching the definition were grouped together. It was our task as terminologists, with the help of subject field specialists, to identify the preferred term to which the others are synonyms e.g. *accelerated weathering* – syn. *artificial ageing*, *applied column* – syn. *engaged column*, *arc boutant* – syn. *flying buttress*, or variants (geographical, orthographical and syntactical variants, acronyms and abbreviations) e.g. *blind alley* – American English *dead-end street*; *block of flats in joint ownership* – American English *condominium*; *downlighting* – Canadian English *pot light*, American English *can light*; *barrel - vault* – variant *barrel vault*; *heating, ventilating and air-conditioning* – abbreviation *HVAC*. Thus, in addition to the entry term and the Romanian equivalent, there are two more data fields (for both languages) that belong to this broader linguistic data category, namely *synonym / sinonim* and *variant / variantă*. A special situation occurred when we could only find an American English term for a concept, while our working language was British English. We decided to solve this problem by writing it down as an entry term, with the specification that it belongs to American English (in brackets, abbreviated to “AE”).

On the other hand, the ideal one-to-one relationship between term and concept becomes very strict when it comes to the terminological entry. That is, in the case of polysemy, when we found a term that had several meanings within our study field we recorded it as a separate entry term, in a separate terminological record, for each concept it was assigned to. Since the Access programme does not permit the occurrence of the same word or expression as entry term more than once, we solved this technical problem by writing a number next to the word, e.g. *drip 1*, *drip 2*.

As the entry term is the most common search item, we recorded it according to the conventional methods of presentation, i.e. in lowercase, in the singular, in the infinitive, etc., unless the plural or the initial capital letter had terminological significance.

3.2. Conceptual information fields

Since terminology is characterized by the conceptual approach, we ordered all the terms we collected by reference to a conceptual system we had previously drawn up. In other words, all the terms in our database were classified according to a structure of knowledge, i.e. narrower subfields of the lexis in the field of architecture and town planning (e.g. architectural theory, architectural elements, construction materials, etc.) e.g. *abbey* – religious architecture, *country house* – residential architecture, *aerated concrete* – architectural engineering / architectural building materials and methods of construction, *facilities management* – administrative and business aspects of architecture. Special emphasis was placed on more recent concepts, such as architectural restoration, landscape architecture and interior design, as well as on traditional Romanian architecture. All these subfields appear in our termbase for both languages under another data category, generically named *domain / domeniu*. This data category is of great importance to special languages users, as it provides a structured comprehensive view of the subject field and better understanding of a concept in relation to neighbouring concepts.

The other conceptual data field is *definition*. It goes without saying that the terminological definition is of utmost importance in any terminographic product, since it links the entry term to the concept which it represents. By tradition, in conceptually-structured databases it is customary to give definitions in one language only (Sager: 1990). We chose English, mainly due to the large number of documentation sources. In many cases we opted for analytical definitions (*genus et differentia*), which systematically identify a concept with reference to the conceptual system and classify the concept within the system, according to its distinctive characteristics. Other types of terminological definitions we used included definitions by paraphrase and mixed definitions, by analysis and description. For each definition the source is given in an abbreviated form, as it

informs on the quality of the documentation and therefore on the validity of the definition. As examples of definitions in the termbase we can mention:

Action architecture 1: architectural style evolving from sketches without precise working-drawings, using materials ready to hand

blind wall: wall without an opening

column base: part of a column which forms the bottom of a vertical supporting pillar

3.4. Pragmatic information fields

There are two data fields providing pragmatic specification of terms in our database, namely *context* (for entry terms) and *usage notes / note de utilizare* (for terms in both languages). As compared to the linguistic and conceptual data fields discussed above, giving examples of the context in which a term occurs is not compulsory in a database. However, following the latest trends in terminology that favour the role of contexts, we decided to insert this data category in our database. Of course, the contexts we provide are explanatory; they bring more information about the concept and are thus complementary to the definition. In addition, contexts show a term in its specific terminological environment in the subject field and in this way assess the validity of the term. Again, the source is mentioned for each context. Examples of contexts for terms of our termbase:

Functionalism " Functionalism grew directly out of the credo that form must reflect function – or 'express' function, as architects like to say." EMA, (Encyclopedia of Modern Architecture) : 112

grisaille glass " This glass gives more light to the interior than the earlier, richer glass work. In grisaille, the plain glass is painted delicately with a floral pattern in grey monochrome." AB (The Architecture of Britain) : 56

Next, the usage notes give information about any special usage of the entry term (e.g. an unusual plural form), as well as make explicit the relationship between various linguistic realizations of the concept (e.g. pragmatic status). Examples of usage notes from our termbase:

adobe - used particularly in Latin America and southwestern USA, adobe produces a distinctive architectural style based on organic forms, a smooth finish, and a minimum of window openings.

art gallery - former term : *art museum*

chamber - now, confined chiefly to elevated style

3.5. Administrative (housekeeping) information fields

This last set of information comprises five data fields, namely *record number*, *author of record*, *date of record*, *author of update* and *date of update*. Given that all terminological information has limited validity in time, these categories are particularly useful with regard to the quality of data; they are certain proof that information in our database can be easily assessed, re-assessed and updated at any moment.

4. Conclusions

At present, "Bază de date terminologice din domeniile arhitectură-urbanism-imbile engleză și română" (Terminological Database in the Fields of Architecture and Town Planning – English and Romanian) contains more than 2000 terms that designate concepts related to various domains in the field of architecture and town-planning.

From among the conclusions, we would like to mention several issues that reveal the way of accomplishing the objectives stated initially in the present paper:

- the attempt to include terms of the most recent subfields of architecture and town planning;

- the creation of a flexible open-ended termbase, built on a rigorous conceptual system that includes the key concepts of the field, representing the starting point for further development (amplification);
- the creation of a customized terminographic product tailored to the specific needs of the Romanian users;
- the general standard of project management applied to all the stages of the project;
- the achievement of proper total quality management throughout the project, based on permanent feedback for previous work and work in progress, based on the principles and methods of modern terminography.

The way in which we carried out the management of the information in our database, as presented above, meets the needs of a fairly wide range of users and serves multiple purposes. We strongly believe that our database is a reliable terminological tool which enhances specialized communication in the fields of architecture and town planning.

References

1. Cabré, M.T., *La terminologie. Théorie, méthode et applications*, Ottawa: Armand Colin – Presses de l'Université d'Ottawa, 1998.
2. Ciobanu, Georgeta, *Îndrumător de terminologie engleză pentru specialitatea arhitectură-construcții*, Timișoara : Institutul Politehnic Timișoara, 1978.
3. ISO 12620 : 1998. *Computer Applications in Terminology – Data Categories*. Geneva: International Organization for Standardization.
4. Pavel, S. and Nolet, D., *Handbook of Terminology*, Canada : Terminology and Standardization – Translation Bureau, 2001.
5. Picht, H. and Draskau, J., *Terminology: An Introduction*, Guildford, Surrey, 1985.
6. Sager, J., *A Practical Course in Terminology Processing*, Amsterdam / Philadelphia: John Benjamins, 1990.
7. Teodorescu, Magda, *Language Practice for Architects to Be*, 2nd ed., București, Editura Universitară "Ion Mincu", 2003; <http://www.iaim.ro/>, 2007.
8. S.E.Wright and G.Budin (eds.) *Handbook of Terminology Management*, Amsterdam /Philadelphia : John Benjamins Publishing Company, 1997.