Alexiev, Boyan

Specialized dictionaries for learners : a Canadian perspective

In: *Variations on community: the Canadian space.* Otrísalová, Lucia (editor); Martonyi, Éva (editor). 1st edition Brno: Masaryk University, 2013, pp. 13-19

ISBN 978-80-210-6404-1

Boyan Alexiev
University of Architecture, Civil Engineering and Geodesy, Sofia, Bulgaria

Specialized Dictionaries for Learners:
a Canadian Perspective

Abstract

This paper discusses the contributions of two leading Canadian terminologists to the development of pedagogically-oriented terminography. Marie-Claude L'Homme proposes a method for converting a terminological database into a specialized learner's dictionary by representing actants in a more user-friendly way, justifying the importance of the latter in describing terms. Semantic roles are represented together with typical terms which instantiate actants in a specialized domain.
Lynne Bowker points out the relevance of using text corpora in the compilation of specialized dictionaries. Meaning is better explained, examples of usage are provided, syntactic behavior is described, and collocations are covered. In general, the corpus-based approach to specialized dictionary making helps in specifying the relative frequency of terms, constructing definitions and identifying phraseological units.

Résumé

Dans le présent article on discute les contributions de deux terminologues canadiennes d'avant-garde pour le développement de la terminographie pédagogiquement orientée. Marie-Claude L'Homme propose une méthode pour la conversion d'une base de données terminologique en un dictionnaire d'études spécialisé par la représentation des actants d'une maniére conviviale, justifiant leur importance lors de la description des termes. On représente des rôles sémantiques de pair avec des termes typiques illustrant les actants dans un domaine spécialisé.
Lynne Bowker met en valeur la pertinence de l'utilisation de corpus textuels dans l'élaboration de dictionnaires spécialisés. La sémantique est mieux expliquée, on donne des exemples d'emploi, on décrit le comportement syntaxique et on y inclut des collocations. En principe l'approche basée sur corpus vers la lexicographie spécialisée aide à la détermination de la fréquence relative des termes, ainsi qu'à la rédaction des définitions et l'identification des unités phraséologiques.

Introduction

The motivation for exploring the issue related to specialized dictionaries designed for learners of both the special language used in a particular discourse community and the special terminology serving as the main building block of that language is the publication of a book by Fuertes-Olivera entitled *Specialized Dictionaries for Learners*. The book features the contributions of two prominent Canadian terminologists to the topic. Their relevance for the development of modern terminological practice is conditioned by the fact that the Canadian School of Terminology nowadays has a leading role in modern terminology science and research. Spe-

cialized dictionaries, resulting from the terminological activities performed within a special domain, have become indispensable tools in today's information and knowledge society. My personal interest in that type of terminological collections has been manifested in a proposal I have made for what I call a ***learner's glossary*** (see Alexiev, ***Knowledge*** and "Terminology") based on the Canadian experience in developing user-oriented terminographic tools.

The aim of the paper is to sum up the views of two Canadian terminologists on designing terminological dictionaries for learners and the application of corpus linguistics in that activity. I will first present briefly the basic contributions of the Canadian School of Terminology. Then, based on the assumptions of the authors regarding the terminological activity in question, I will discuss a proposal for designing terminological dictionaries for learners based on a lexical semantic approach to representing actants and assess the role of Corpus Linguistics in developing specialized learners' dictionaries.

Canadian contributions to terminology science and practice

The contributions of the Canadian School of Terminology can be summarized in the two main approaches to terminology, followed by terminologists from two Canadian universities. It is important to emphasize that despite the differences in the methodology they propose for processing terminological data, they both rely on computerized textual corpora.

The first approach can be defined as a ***conceptual corpus-based approach*** (L'Homme, "A Look") followed by the terminology research group at the University of Ottawa led by Prof. Lynne Bowker, Director of the School of Translation and Interpretation. The approach can be generally defined by the following characteristics:

1. It is basically onomasiological, i.e. concept-based. In other words, the terminological data processing proceeds from the concept to the term designation (linguistic expression)
2. It proposes a new form of terminological database called 'terminological knowledge base' or TKB.
3. A TKB organizes concepts into networks of relations including both hierarchical (generic–specific and whole–part) and non-hierarchical (cause–effect and object–function).
4. The approach involves work on identifying conceptual relations in running text.

The second approach is known as a ***lexico-semantic corpus-based approach*** followed by the terminology group at the University of Montreal led by Prof. Marie-Claude L'Homme, Director of OLST. The main characteristics of that approach can be summarized as follows:

1. It is semasiological, i.e. lexically-based, which means that terminological data processing proceeds from the term designation (linguistic expression) to the respective concept.
2. The theoretical basis of the approach is Explanatory and Combinatorial Lexicology (Wanner).
3. An emphasis is laid on the linguistic properties of terms such as collocational behaviour, semantic relations, synonymy, antonymy, argumental relationships/ actantial structures, etc.

Designing terminological dictionaries for learners based on Lexical Semantics

L'Homme makes a proposal for converting a terminological database containing French terms (Computing and the Internet) encoded according to formal lexical semantics (principles of Explanatory and Combinatorial Lexicology) into a specialized learners' dictionary ("Designing"). The proposal is based on the assumption that actants can be represented in specialized terminological collections, thus providing additional actantial information. An increasing number of dictionaries (general-language or specialised) represent actants explicitly or implicitly in articles. Actants are defined as the obligatory linguistic participants involved in the meaning of predicative linguistic units (verbs and adjectives), but also de-verbal and de-adjectival nouns. For example:

LOAD: X *(someone)* loads Y *(something)* into Z *(something)*
X *(someone)* loads Z *(something)* with Y *(something)*

The explicit representation of actants is a recent addition to various types of dictionaries but is more common in learners' dictionaries. For example, in the ***Collins Cobuild English Language Dictionary*** the representation of actantial information is realized by including actants in the definitions:

EAT: when ***you*** eat ***something***, or when ***you*** eat, ***you*** put ***food*** into your mouth...

As can be seen from the example, the actants are represented by pronouns or typical lexical units. In line with the descriptive principles of the Explanatory and Combinatorial Lexicology, actants play a central role in the description of the sense of the linguistic units. They are stated in the actantial structure at the beginning of the entry and labelled in terms of actantial roles. Role labels are designed to represent the semantic relationship between the predicate and its actants. The theoretical basis for representing the actantial structure is Fillmore's Frame Semantics (Fillmore). For example:

Print: Agent ~ s Patient with Instrument

In the original database, ***DiCoInfo*** (Le dictionnaire fondamental de l'informatique et de l'Internet), the actants are stated in definitions. For example:

Configure: Agent ~ Destination → Def: An Agent sets the parameters of a Patient so it can work according to those of specific hardware or software

Actants are also necessary to explain most lexical relationships (especially syntagmatic).

E.g. ***space bar***: ~ used by Agent to act on Patient

It should be noted that actantial structure is not a data category commonly added to dictionaries. In principle, the representation of actants in terms of actantial roles is rather opaque to most users. Most explanations of lexical relationships are also opaque since the first level of explanation refers to the actantial roles and the second level uses lexical functions (e.g. *InceptReal 1*) proposed in the Explanatory and Combinatorial Lexicology.

L'Homme proposes a solution to this problem. The option is a system in which actants are represented by typical terms instantiating actants in the respective specialized field. Below is an example of the actantial structure of the computer term ***click*** accompanied in brackets by typical terms which represent the actants:

Click: Agent (user) ~ on Patient (file, icon) with Instrument (mouse)

Therefore, the typical terms appear in the definition and the explanations given for lexical relationships. It is also important to point out that the two representation systems are maintained in the database. The advantages of using typical terms can be summarized as follows:

(a) Since typical terms instantiate actants in concrete contexts, users can access information on the combinatorial possibilities of the lexical unit.
(b) If typical terms are chosen wisely, users will infer the correct information related to the actants, e.g. the actantial structure for ***click*** (humans play the role of agents, any icon/file can be clicked).
(c) Typical terms correspond to terms in the respective domain, so they are expected to be described elsewhere in the dictionary;

Despite the advantages mentioned above, the choice of the ideal typical term often poses a challenge. There are four criteria for a term to be defined as a typical term, namely, (1) it should be a term belonging to the respective special domain; (2) it should have a high frequency of occurrence in the corpora; (3) it should be a generic term; (4) in the explanations of lexical relationships or in the definition, a typical term should be the most "natural" one to appear in this context. Sometimes these four criteria cannot be applied simultaneously**.** For example, the patient role in ***click*** cannot be exemplified by any single generic term. Hence, the solution to this problem is to represent it by two typical terms: ***file*** and ***icon***.

The contribution of Corpus Linguistics to the development of specialized dictionaries for learners

Lynne Bowker examines how corpus linguistics has impacted both the process and product of specialized lexicography, with a particular focus on specialized dictionaries for learners. To understand the meaning of the latter concept, it is necessary to define the two subordinate concepts, viz. specialized (also known as terminological or technical) dictionaries and learner's dictionaries.

Specialized dictionaries are dictionaries that (a) treat specialized fields of knowledge (e.g. business, chemistry, law); (b) focus on LSP, i.e. consist of lexical items (terms) used to describe

concepts in a special domain; (c) can be maximizing (covering a large portion of the respective terminology or minimizing (covering a limited portion of that terminology); (d) can be monolingual or bilingual; (e) have content and design determined by the needs of the intended users (experts, semi-experts, laypersons). Regarding encyclopedic knowledge, specialized dictionaries provide less information for experts and more specialized knowledge for semi-experts and laypersons. With respect to linguistic knowledge, if the users are non-experts, they are likely to require more information on, e.g. collocations, irregular inflections, etc.

A learner's dictionary can be defined as a general monolingual dictionary specially conceived for non-native speakers of a language. It is traditionally compiled with the decoding learner in mind. Nowadays learner's dictionaries are used not only by young foreign language users but also by mature users for whom the foreign language has become their professional ***lingua franca***. In other words, they need contextualized dictionary information that can help them comprehend and produce adequate texts.

Therefore, a need has arisen for dictionaries intended to help users learn about the concepts and terms in a specific field of knowledge in one or more languages with specific contents determined by the intended user group and purpose of the dictionary. For example, a monolingual or bilingual production- or reception-oriented dictionary may target subject field experts who are language learners, or language professionals who need to learn about the subject field and its associated LSP. To meet this need, new types of 'hybrid' learner's dictionaries have appeared which combine both general and special knowledge. A good example is ***Cambridge Advanced Learner's Dictionary Business Plus*** (2009), based on the 100-million-word Cambridge Corpus of Business English which "offers everything you would find in a general English dictionary PLUS business-specific vocabulary." Another example is the corpus-based ***Macmillan English Dictionary*** (MED). A survey conducted among nearly 2,000 users of that dictionary showed that a growing number of language learners currently using English when studying or working in fields such as medicine, business, or information technology expect the dictionary in question to explain the special terms they encounter on a regular basis. The modern approach to studying a foreign language presupposes that learners should study authentic examples of what people have actually written or said, rather than hypothesizing about what they might or should say. Hence, the reference tools helping those learners, i.e. the dictionaries, should be compiled based on large corpora of natural language.

The concept ***corpus*** can be generally defined as a sample of language systematically gathered for a particular purpose and coherently organized for that purpose. In its modern context a corpus is a relatively large collection of authentic texts gathered in electronic (capable of being processed by a computer) form according to a specific set of criteria (general or special texts, written or spoken, etc.). For example, ***Collins Cobuild English Language Dictionary*** (1987) was the first dictionary based on a systematic analysis of a large-scale corpus. The use of corpus data has transformed not only the process but also the product. The dictionaries designed for learners have improved enormously. Along with the traditional functions of such dictionaries, namely explaining meanings, giving examples of usage and describing syntactic behaviour, these reference tools nowadays provide more systematic coverage of pragmatics, register and collocation. The corpus-based approach to general and specialized lexicography has contributed significantly to improving the dictionary macrostructure. The frequency data

from a corpus helps to inform the selection of lexical items to be included as headwords and to construct definitions composed of those words. The different meanings of a word are listed in an order based on the frequency of occurrence in corpus. The latter provides lexical, grammatical and pragmatic information including authentic examples, grammar behavior (e.g. whether a noun is typically used in the plural), collocational behavior, style, register, genre, etc.

Concerning specialized learners' dictionaries, it should be noted that the use of corpus-based methods has been slower to take hold in specialized lexicography than in general lexicography. One reason for that is the insufficient amount of online technical texts as compared with general language ones. However, the situation is rapidly changing with the advent of high technology and the consequent digitalization of our social and cultural life resulting in a considerable increase in the amounts of online specialized texts.

Bowker suggests that a corpus-based approach could be applied to the creation of specialized dictionaries for learners in the same ways already applied to general learner's dictionaries. It can help dictionary compilers in identifying the relative frequency of terms, in constructing definitions, selecting authentic examples, identifying phraseology, etc. The 'hybrid' approach to combining general language words and specialized terms in specialized dictionaries for learners, for example, based on the corpus frequency data, can help to identify only those general language words which are used and recur most often in the respective special language. Learner corpora could be used to provide help for those users with encoding needs. Bowker concludes that "the adoption of a corpus-based approach in specialized pedagogical lexicography seems to be on the horizon" (168).

Conclusion

It is interesting to note that under the strong influence of the Canadian School of Terminology, when proposing my own small-scale terminological collection, viz. a learner's glossary (Alexiev, ***Knowledge*** and "Terminology"), I followed somehow intuitively the main assumptions of the two Canadian terminologists discussed above with regard to using text corpora and representing actants instantiated with typical terms. When collecting my study corpus on the construction material ***concrete*** I resorted to three main genres in the domain: a recent text on ***concrete*** from a reliable university textbook (didactical), an online primer (basic didactical) and a glossary of ***concrete*** terms. The frequency of occurrence was used to determine the head terms belonging to the limited subdomain of ***Concrete***. I developed a classification scheme of the general aspects of the basic concept/top term: ***Concrete***: types, composition, properties, technology and use. On that basis, I represented the actants in a user-friendly manner, thus making the glossary easily comprehensible for the target group of users, i.e. my architecture and civil engineering students and translators in the Civil Engineering domain. Below are examples of contexts to be used in the glossary entry:

Concrete (object) ***hardens*** in formworks (location).

Builder (agent) ***places*** concrete (object) in formworks (location).

Therefore, from my personal experience as a terminologist, I can conclude that the proposals made by the two Canadian terminologists discussed above could be successfully applied not only in compiling terminological collections for learners but also in teaching LSP and conducting technical translation courses.

Works cited

Alexiev, Boyan. *Knowledge-Oriented Terminography*. Sofia: Avangard Prima, 2011. Print.

----------. "Terminology structuring for learner's glossaries." *Knowledge Organization*, 33.2 (2006): 96–118. Print.

Bowker, Lynne "The Contribution of Corpus Linguistics to the Development of Specialized Dictionaries for Learners." *Specialized Dictionaries for Learners*. Ed. Pedro A. Fuertes-Olivera. Berlin/New York: De Gruyter, 2010. 155–168. Print.

Fillmore, Charles J. "Frame semantics." *Cognitive Linguistics: Basic Readings* 34 (2006): 373–400. Print.

Fuertes-Olivera, Pedro A., ed. *Specialized Dictionaries for Learners*. Berlin/New York: de Gruyter, 2010. Print.

L'Homme, Marie-Claude. "Designing Terminological Dictionaries for Learners based on Lexical Semantics: The representation of actants." *Specialized Dictionaries for Learners*. Ed. Pedro A. Fuertes-Olivera. Berlin/New York: De Gruyter, 2010. 141–153. Print.

----------. "A Look at Some Canadian Contributions to Terminology." *Modern Approaches to Terminological Theories and Applications*. *Linguistic Insights*. Ed. Heribert Picht. Bern: Peter Lang, 2006. 55–75. Print.

Wanner, Leo, ed. *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam & Philadelphia: Benjamins Academic Publishers, 1996. Print.