

The 8th International Scientific Conference
eLearning and software for Education
Bucharest, April 26-27, 2012
10.5682/2066-026X-12-017

**TEACHING E-COMMERCE TERMINOLOGY BY CREATING ELECTRONIC
CORPORA**

Mariana COANCĂ
Romanian-American University, Bucharest, Romania
E-mail: coancamaria@yahoo.com

Abstract: *In this paper we emphasize the fact that teaching e-commerce terminology is not an easy-doing task, taking into account that terminology is an interdisciplinary science that deals with the classification and standardisation of the specialised vocabulary, by giving relevant information in a technical-scientific framework. The major goal of this science is to enforce the terms and describe them according to proper and current communicative situations, using electronic corpora and then applying the methods of linguistics (the linguistic description of terms). Nowadays, teaching new, emerging terminologies by creating an electronic corpus has been strengthened at an international level. Teaching e-commerce terminology can be very challenging for teachers and students: on the one hand, it cannot be achieved without the linguistic analysis of the terms, but on the other hand students become more and more interested in the new terminological trends base don the creation of electronic corpora. Furthermore, e-commerce terminology becomes functional if the teaching process is also based on a tight collaboration between teachers and e-commerce experts, helping teachers to present and explain e-commerce terms in a vulgarised manner. Teachers, in their turn, manage to find suitable teaching methods, after the creation of the electronic corpus in order to enhance the students' knowledge in an interdisciplinary approach. Thus, teaching is productive after the linguistic analysis of e-commerce terms and the functional degree of certain terms is relevant in various fields: economics, marketing, information technology.*

Keywords: *e-commerce terminology, teaching, electronic corpus, WebBootcat*

**I. THE DOCUMENTATION PROCESS AND THE ELECTRONIC CORPORA
USAGE**

The main mediator of this terminology at international level is the American English, having consequences upon the expression and formation of e-commerce terms. Thus, English can be considered lingua franca of the business environment, in which the major role is played by communication for establishing and developing *professional relations*.

Another reason that justifies the frequent use of the Internet resources, is represented by a *strong interdisciplinarity of the e-commerce field*, consisting of a mixture of terms and concepts which belong to other disciplines or different fields. This peculiarity requires the documentation in certain fields related to *economics* and *information technology* (marketing, criptography, commercial transactions, etc.), the Internet being a very important and reliable source of documentation to make the needed research. Furthermore, the steady development of the specialized languages is another reason to use the Internet resources, as a fundamental source of documentation and we take into consideration that among the essential characteristics of this method one can notice the continuous updating of the online available materials.

The delimitation and the defining of the corpus are two important stages, but we also focus on the contents of our corpus to ease the teaching and learning process: types of online texts, manuals,

magazines, taking into account a corpus is limited by time and space. In our case, this aspect is noticed because we use web pages which contain *electronic texts*. Obviously, we can say the press language is quite sensitive to the lexical fluctuations: the press is defined according to the latest news and this aspect conditions the approached topics. Therefore, regarding the teaching and learning of e-commerce terms, it is highly important for us to justify the choice of the corpus, of the electronic texts, of the investigated period. In our opinion, there is a relation among the Internet dynamics, the e-commerce dynamics and the linguistic dynamics of the e-commerce field in our country; for this reason, the latest years are a relevant specimen for the dynamics of the e-commerce field and the selection of electronic texts is very useful to the creation of an electronic corpus, from which the students can extract and analyse the most frequent terms of e-commerce.

The creation of a corpus depends to a large extent on the type of research one is going to undertake, just like the results of the research depend on the creation of the corpus. If the terminological activity consider the electronic texts as the natural environment of the *terms*, the lexicography is essentially normative, because this approach is not always according to the norms. Certain dictionaries based on language are frequently drawn upon the literary texts and not drawn upon the spoken language. Thus, the language of dictionaries do not show how the language must be spoken, but how the lexicography considers it must be spoken (Cabr  1999: 37-38). This method can be considered the opposite of the teaching and learning according to the investigated corpora. The use of traditional corpora and mainly electronic corpora, allows the collecting of a lot of information and it represents a reliable method to create online dictionaries and glossaries.

The corpus can deliver contexts which show the meaning and the use of the terms. Furthermore, the context allows us to notice the meaning changes and to analyse the vocabulary evolution (the use of new words, the disappearance of words or meanings, etc). If it is permanently updated, the corpus is a very good „observer” of a specialized language, in our case the e-commerce language. The documentation is a self-stagnating activity within the terminological activity. Some authors consider that the “terminological units are included in the *specialized languages* and they behave naturally in the *specialized languages*” (Cabr  1999: 80). It is unquestionable the fact that, talking about specialized languages, it is not debated the „legitimacy” of the text, its existence according to the dictionary. If we ask ourselves where we can find the specialized terms, the logical answer is: in the specialized texts, in the specialists’ speeches. If the startup of terminology was strictly onomasiological, the development of *corpus linguistics* made the terminologists approach the semasiology, the same manner it made the lexicographer focus on the text-based research, on the peculiarities of a language specific to a specialized field.

II. ELECTRONIC CORPORA AS A METHOD OF TEACHING AND LEARNING THE E-COMMERCE TERMINOLOGY

The creation of an *electronic corpus* was an essential stage for teaching and learning the e-commerce terminology, including the linguistic analysis of the e-commerce terminological units. According to Sinclair (1991: 171) a corpus created adequately, offers information about the authentic use of the terms specific to a certain field, contributing to the solving of lexical, phraseologic and syntactic issues arising in the specialized translation, too. While creating the corpora, we tried to meet the requirements of modern linguistics, which are indispensable for the creation of corpora: *representativeness, finite size, reference standard*. Regarding the first requirement, the corpus must be a proper representation of the linguistic variety and it must show the real use of the e-commerce language. In order to ensure this aspect, one can start from a nucleus consisting of documents, that deal with the respective topics in a certain field or through a series of linguistic instances (see also <http://www.realiter.net/spip.php?rubrique38>).

The second characteristic of electronic corpora is, as previously mentioned, the finite number of electronic texts they contain. This limitation will be made through a careful selection of the texts that are significant from the quantitative point of view, thus meeting the requirement based on representativeness mentioned above. We would like to mention that the second feature of corpora is not static at all and the requirements for understanding the e-commerce terminology, as well as the need to reach equivalent terms has led to the enrichment of the initial nucleus of texts, because they prove to be very useful, according to the assertions of some authors: “a corpus (...) must be ample” (Sinclair 1991: 18). This author speaks about *trimming*, by “adding extra texts where there are gaps”, in order to achieve the goals of the research (ibidem: 20). Regarding the dimensions of the corpus, the same author suggests that these corpora should consist of entire texts, not of simple fragments. This solution prevails, because certain linguistic peculiarities, that must be noticed in the text, are not evenly distributed in the text. We took into account this proposal, by keeping the original contents of the texts and their original format.

The third characteristic of corpora is represented by their electronic format, which enables qualitative and quantitative investigations, that are harder to be achieved by traditional methods of investigation, thus giving information on the spot, regarding the use of the e-commerce terms. Then it can be used as a starting point for future research. This corpus can be useful to the teaching and learning process, because students carry out linguistic analyses on the the extracted terms and make comparisons between the e-commerce terminology used in a language or two languages. Besides the already mentioned peculiarities, while creating the *electronic corpus* we also took into account other aspects that improve the quality of the electronic texts, meeting the representativeness requirement for the linguistic analysis of the e-commerce terminology, according to the principles of teaching and learning a specialized terminology.

The choice of electronic texts was made, taking into account their origin, giving a full priority to the electronic texts belonging to certain institutions or authors whose articles deal with issues about the e-commerce field (the articles were downloaded from the *Epayment blog*, the electronic magazines *Link2eCommerce* and *Underclick*). Another requirement is based on the fact that the choice of electronic texts must be chronological, we especially chose recent, updated electronic texts that mirror in an adequate manner the rapid linguistic changes that characterize the e-commerce field. Another significant criterion is based on the addressee’s regard on the electronic texts of the *corpus*. In this case, it was given a greater importance to the texts for the specialists in e-commerce, but certain texts target the non-specialized public, that is highly interested in the presented topics.

III. THE SPECIALIZED CORPUS OF E-COMMERCE

Bowker and Pearson alleged that manuals are not suitable to observe specialized terms, because they use a vulgarised language excluding the complex, specialized designations. For this reason students are instructed to permanently check if the selected terms are accepted and used by the specialists in the respective field, because the linguistic peculiarities involve even more documentation. The students could be also advised by their teacher to visit forums and find terms specific to the e-commerce field, but sometimes they run into difficulties because the discourse on the respective sites could be quite familiar and the discussions could contain abbreviations of the terms that are not understood by the Internet surfers. Unlike other fields, such as medicine, biology, it is difficult to identify all the specialists of the Internet field. Those connected to the Internet network have heterogeneous goals and preoccupations – the creation of programs, websites and the assurance of data security, etc.

The nature of the web pages and the electronic texts we selected is in accordance with the students’ goals, because they wanted a corpus containing e-commerce terms that are used by

specialists and the non-specialized people. We would like to add that the electronic dictionaries are useful, although some authors mentioned they do not register the entire specialized language. In a general dictionary one cannot find terms used in the technical articles, simply because the dictionaries are not updated. A few glossaries and dictionaries are dedicated to the Internet and the e-commerce field and a few e-commerce terms are registered by the IT and marketing dictionaries. Thus, in our opinion there should be a steady collaboration between the specialists in e-commerce and terminologists to provide adequate definitions for the e-commerce terms, to elaborate electronic dictionaries and glossaries in order to meet the students' requirements, because their further activity focuses on identifying the equivalent terms for the English e-commerce terms.

Therefore, the electronic corpus, that is used for pedagogic and learning purposes (for instance, the extraction of e-commerce terms, their translation in the target language and their linguistic analysis based on their compounding methods and semantic relations, etc) can be achieved with various software programs, but we opted for the facility offered by the program called WebBootCat, as a method of building an electronic corpus, involving the search engine Google, too. The World Wide Web is divided in notional fields, that help us to build a corpus consisting of web pages, presented either in the native language or in other foreign languages. The program WebBootCat is very useful to linguists, translators, students and researchers that participate in research projects and are very active in the linguistic and translation studies. This application is available within the Terminology Laboratory, SSLMIT Forli, University of Bologna. According to the instructions, we inserted in the first box a few *keywords* in inverted commas (these keywords can be monosyllabic or plurisyllabic), consisting of e-commerce terms, also called "seed terms" (for instance, *e-commerce, online store, online payment, site*). In the second box, we enclose the API Google key, which must be obtained according to the Google requirements (therefore, an application is filled out to make an account and obtain the API Google key). In the third box one can choose the language and in the last boxes we write the name of the corpus, respectively the e-mail address of the user so that he could be announced when the corpus is built (Baroni et al., 2006: 248-249).

This procedure did not take us too long and we built an electronic corpus, consisting of web pages that correspond to our study (texts from the *Epayment blog*, the electronic magazine *Link2eCommerce*, *Underclick*. To improve the teaching process we filtered the most frequent e-commerce terms and afterwards we established the peculiarities of the selected terms, according to the requirements of a linguistic analysis. Furthermore, the tight collaboration between students and teachers can lead to the creation of a terminological database, according to various criteria (Pearson 1998; Scarpa 2001; Magris et al. 2002; Osimo 2000). Students are acknowledged that it is highly important for terminologists and especially for translators to know how to recognize the most frequent terms; because obviously a total monosemantic correspondence between the concept and the term expressed in different languages will not occur. Moreover, Scarpa (2001: 155) mentions that "the terminological units are used in the specialized context or in the standardized form".

This corpus proved to be user-friendly, offering students the possibility to find new terms specific to the e-commerce and they can also extract definitions of the identified terms from the electronic texts (Kocourek 1991: 23). This aspect is highly important to the technical disciplines, in which the terms designate precise concepts, the students' task being to identify the terms, the English loans, the acronyms specific to e-commerce terminology. Their linguistic analysis on the created corpus proved to be very productive: from a quantitative point of view they extracted a lot of fixed syntagms, in which the particle *online* or *electronic* is valuable because it establishes the e-commerce meaning (the following syntagms have a high frequency in the electronic corpus: ONLINE SHOP, ONLINE SHOPPER, ONLINE SALE, ONLINE RETAILING, ONLINE PURCHASE, ONLINE INVOICING, ONLINE CONSUMPTION, ONLINE CATALOGUE, ONLINE AUTHORIZATION, ONLINE AUCTION, ONLINE ADVERTISING, ELECTRONIC PAYMENT, ELECTRONIC MONEY, ELECTRONIC STORE, etc). The presentation herein shows the interface of the WebBootCat application, the search page, that was downloaded from the online article *WebBootCaT: a web tool for instant corpora* (authors Marco Baroni, Adam Kilgarriff, Jan Pomikálek and Pavel Rychlý, 2006: 247/252):

Firefox

marks Tools Help

http://corpora.fi.muni.cz/bootcat/

Search PageRank ABC Check AutoLink AutoFill Options

alpha

BootCaT

www version

Seed words

Enclose multiwords expressions into quotes (").

Your Google key

This application uses Google API. You need a Google key to be able to use it. If you do not have one please go to <https://www.google.com/accounts/NewAccount>.

Language

Select the language of the corpus to be built.

Select URLs

Check this box if you would like to manually filter found URLs before the webpages will be downloaded.

Tag corpus

Your corpus will be POS-tagged and lemmatised using the [TreeTagger](#). Following languages are currently supported: English, French, German, Italian, Spanish. This options has no effect if used with any other languages.

Corpus name

Chose a name for your corpus.

Your email address

The time needed for building a corpus is highly variable and may take minutes, or hours. If you enter your email address you will be notified when the corpus is ready to use.

IV. CONCLUSIONS

In conclusion we can say that the corpus is an essential tool in:

- *the acquisition stage*: the e-commerce terms are selected.
- *the representation stage*, through which the linguistic peculiarities of the selected e-commerce terms are highlighted..
- *the explanation stage*, in which the corpus is accepted and students extract the context of the term usage, in order to offer a definition for the e-commerce terms.

The process of terminological systematization through a corpus building is an opportunity to promote a theoretical idea about the essential peculiarities of the e-commerce language, in comparison with the peculiarities of the specialized languages in general. Therefore, the students, assisted by us, carried out a linguistic investigation that reflects the following aspects: the permanent use of the new designations (compound forms, acronyms), the great impact of the English language upon the Romanian language, the steady use of the English terms or the use of their Romanian equivalents. Moreover, the students noticed that many English terms are accepted, but the English loans specific to the e-commerce field are mainly used by the specialists in e-commerce. Students used the electronic

texts and searched keywords and this approaches triggered the active participation of the students in the linguistic investigation.

Teaching the terminology of a specialized field could be very challenging for teachers and students and the building of an electronic corpus improves the linguistic investigation for the specialized terms, on the one hand and promotes the most frequent terms, on the other hand. Last but not least, students can elaborate multilingual glossaries of specialized terms for an effective translation process of specialized documents.

References

- [1] Aston, G., 2000, "I corpora come risorsa per la traduzione e per l'apprendimento", in S. Bernardini e F. Zanettin (a cura di) *I corpora nella didattica della traduzione*, Bologna: CLUEB, pg: 21-29.
- [2] Baroni, M., Kilgarriff, A., Pomikálek, J., Rychlý, P., "WebBootCaT: instant domain-specific corpora to support human translators", Proceedings of EAMT 2006, Oslo, 2006, pg: 247-252.
- [3] Bessé (de) Bruno, 1994, "Le domaine", in *Français scientifique*, pg.182 and the following.
- [4] Bidu-Vrănceanu, Angela, 2007, *Lexicul specializat în mișcare. De la dicționare la texte*, Editura Universității din București.
- [5] Bidu-Vrănceanu, Angela, Alice Toma, Roxana Ciolăneanu, Elena Museanu, 2009, "Terminologia economică internațională – probleme ale comunicării plurilingve", comunicare la Zilele Realiter, 2009, <<http://www.realiter.net/spip.php?article1762>>.
- [6] Bidu-Vrănceanu, Angela (coord.) 2010, *Terminologie și terminologii*, Editura Universității din Bucharest.
- [7] Bowker, L., 1998, "Exploitation de corpus pour la recherche terminologique ponctuelle", in Humbley John, *Terminotique et documentation*, Terminologies nouvelles, no. 18, June 1998, pg: 22-27.
- [8] Bowker, L., 1998, "Using specialized monolingual native-language corpora as a translation resource: a pilot study", *Meta*, 43/4, pg: 631-651.
- [9] Bowker, L., J. Pearson, 2002, *Working with Specialized Language: A Practical Guide to Using Corpora*, Routledge, New York, October.
- [10] Cabré, Maria, Teresa, 1996, "Terminology today", in *Terminology, LSP and Translation*. H. Somers. Amsterdam, Benjamins, pg: 15-34.
- [11] Cabré, Maria, Teresa, 1999, *Terminology: Theory, methods and applications*. Amsterdam, Philadelphia: John Benjamins Publishing Company, available online www.library.nu.
- [12] Kocourek, R., 1991, *La langue française de la technique et de la science. Vers une linguistique de la langue savante. 2ème Édition*. Wiesbaden : Oscar Brandstetter Verlag.
- [13] Magris, M., 2002, "Le relazioni concettuali", in Magris, M., M.T. Musacchio, L. Rega & F. Scarpa, *Manuale di terminologia. Aspetti teorici, metodologici e applicativi*. Milano: Hoepli. 151-166.
- [14] Museanu, Elena, 2011 *Terminologia economică. Caracteristici în româna actuală*, Editura Universitară
- [15] Osimo, B., 2000, *Traduzione e nuove tecnologie. Informatica e Internet per traduttori*. Milano: Hoepli.
- [16] Sager, J., 2000 – "Pour une approche fonctionnelle de la terminologie", in H. Béjoint et Ph. Thoiron, pg. 40 and the following.
- [17] Scarpa, F., 2001, *La traduzione specializzata*. Milano: Hoepli.
- [18] Sinclair, J.,M., 1991, *Corpus, Concordance and Collocation*. Oxford : Oxford University Press.

ELECTRONIC SOURCES:

The online press release Pyramid Grup Legal
<http://www.realiter.net/spip.php?rubrique38>