Mafalda ANTUNES (mca@iltec.pt)
Maria DORIA (mad@iltec.pt)
FCT[1] / ILTEC / A*i*T

The *Dicionário de termos do comércio electrónico /* Dictionary of e-commerce terms (DTCE): conception and methodological options

1. Introduction

The main goal of this paper is to present the project of conception and elaboration of the *Dicionário de Termos do Comércio Electrónico* (DTCE). The DTCE is a project from ILTEC[2] – Instituto de Linguística Teórica e Computacional (Institute of Theoretical and Computational Linguistics (www.iltec.pt)) and from ahp – Application & Hosting Provider (www.ahp.pt) that aims to produce a dictionary that includes terms in Portuguese and its equivalents in English, for further edition in electronic support (CD-Rom and/or Web).

The production of a dictionary with such characteristics raises some problems essentially related with the fact that e-commerce is a recent specialized domain with an unstable terminology, and a strong presence of English terms, therefore needing an urgent terminological intervention.

In this paper we intend first to present some of the e-commerce vocabulary characteristics, which justify a dictionary in such domain. Secondly, we intend to present the structure of the DTCE. Finally, given the adopted methodology we will present real examples questioning some problems related to the integration of loan words, the adoption of equivalents and graphical criteria.

2. Characteristics of e-commerce vocabulary

ILTEC, A*i*T – Association of Terminological Information (www.ait.pt) in particular, has collaborated in the *Panlatin Glossary of E-Commerce.* This collaboration consisted in the establishment of equivalents, in Portuguese, for a list of about 200 e-commerce terms. The main goal of this project was the edition of a glossary containing equivalents in the following languages: Catalan, Spanish (Spain and Mexico), French (France and Canada), English (USA and Canada), Portuguese (Brazil and Portugal) and Romanian (Romania and Moldavia).

This project was part of Realiter, coordinated by Gabriel Huard from the *Centre de Terminologie et Normalisation* (Montréal, Canada).

The nomenclature for this glossary was given in English and French which determined the presented solution for some terms. However, this factor was simultaneously unstable in some cases, once we identified some differences in the two languages. The proposed list of terms for Portuguese was the result of a linking of several factors such as the use of trusted sources, the consultation of specialists and when possible the translation, always focusing on the attempt to create true equivalents in Portuguese indeed used as an alternative to English.

Our participation in this project allowed us to realize that e-commerce vocabulary has its own properties that should be the target of a more serious study. For this reason, we

---

[1] Participation of Maria Doria in this congress was funded by the Fundo de Apoio à Comunidade Científica (FACC) from FCT.

[2] Being developed in particular by A*i*T, the terminological group from ILTEC.

have sketched a project of conception and elaboration of a dictionary of e-commerce terms in European Portuguese (PE).

E-commerce is an area spread through the Internet that is nowadays quite largely developed. Its strong impact favours the birth of new terminological units. Consequently, the corpus that constitutes the basis of this dictionary is mainly extracted from the Internet. This new vocabulary reveals great instability once it belongs to a new area in continuous evolution. We consequently find important to collect this vocabulary as well as study its evolution.

Hence, the neological character of the terms that form the DTCE's nomenclature is responsible for some instability that one can verify not only in the way the English terms are adapting to the Portuguese linguistic system, but also in the evolution of the lexical field's structure of e-commerce.

According to the proposal of Antunes M., R. Gonçalves & S. Correia (2002: 122) on "e-terms"[3], the e-commerce terms reveal as well "a great instability reflected at several levels: at the graphical and orthographical level, once the typographical aspects present some variety; at the morpho-syntactical level, once they present some specificity in the way they are formed; finally, at the semantic level, due to the easiness with which these terms are exposed to semantic extensions to other domains of experience (...)" (our translation).

The import of science and technology usually comes with the import of terms belonging to this domain. In the specific case of e-commerce, the vocabulary is established with terms coming mostly from English. For this reason, its form of integration and adaptation to our linguistic system is irregular as we can see in the following examples:

– Borrowed words used in their original form: *e-business, e-commerce, banner.*
– Hybrid constructions: *e-mail comercial, trocas on-line, firewall de segurança.*
– Creation of equivalents respecting the structure of the English basis: *cibercomprador* (*cyber shopper*).
– Creation of equivalents respecting the Portuguese syntax: *negócio-e* (*e-business*), *cartão inteligente* (*smart card*).

3. DTCE's presentation

The DTCE is presently in its conclusion phase and has about 400 terms in European Portuguese within the e-commerce domain, with its definitions and equivalents in English.

Among other things, ahp was responsible for the informatics support in what concerns the construction of the database, the terminological record and the search engine following the guidelines given by ILTEC. The terminological record conceived for this dictionary contains a field referring to the reliability level of the terms which allows us to find the term more acceptable in Portuguese contributing to the terminological harmonization of this specialized domain.

Concerning the macro-structure, besides a brief introduction presenting the dictionary, the DTCE includes the conceptual structure of this domain and the terminological records, each containing one term.

Since this is a dictionary in electronic support that involves two languages (Portuguese and English), we find that a search engine would make its use easier, since it allows the search by "Portuguese term", "English term", "subdomain" and "term(s) included in the definition". Figure 1 presents this search engine:

---

[3] These authors understand "e-terms" as "each one of the units that present in its structure the particle *e* with the meaning *electronic* (examples: *eCommerce* – for *electronic commerce* –, *comércio-e* – for *comércio electrónico*, etc.)" (our translation).

Termo em Português	
Termo em Inglês	
Subdomínio	
Definição	
	Procura

Figure 1 – Search engine.

Referring to the micro-structure, each term was inserted in a terminological record containing the following fields: "Portuguese term"*, "Morpho-syntactic category"*, "Reliability level"*, "Synonymous", "English term"*, "Subdomain"*, "Definition"*, "Definition (source)"*, "Context", "Context (source)", "Notes", "User", "Revision date", "Linguistic revision date", "Final revision date". From these the ones pointed out with an asterisk are obligatory fields.

The "morpho-syntactical category" field allows us to choose between one of the following categories: "s. m." (masculine noun), "s. m. pl." (masculine plural noun), "s. f" (feminine noun), "s. f. pl." (feminine plural noun), "s. com. de 2 sing." (compound noun with 2 sing.), "s. com. de 2 pl." (compound noun with 2 pl.), "adj." (adjective) and "v." (verb). As to the "reliability level" it is also possible to select one of the following options: "termo a ser evitado" (term to be avoided), "proposta para termo" (proposal for term) e "termo bem construído" (well formed term).

Termo em Português		*
Categoria Morfossintáctica	Seleccionar Opção ▾	*
Índice de Fiabilidade	Seleccionar Opção ▾	*
Sinónimos		
Termo em Inglês		*
Subdomínio		*
Definição		*
Definição (Fonte)		*
Contexto		
Contexto (Fonte)		
Observações		
Utilizador		
Data Revisão		
Data Revisão Técnica		
Data Revisão Linguística		
Data Última Revisão		
	Inserir	

Figure 2 – Terminological record.

4. Methodology

Following Cabré's point of view (1999: 115) we believe that "terminography involves gathering systematising, and presenting terms from a specific branch of knowledge or human activity. (...) Even though it is practical in nature, terminography is not an independent activity that individual specialists can deal with on their own but is governed by a series of technical, formal, and procedural recommendations that have been internationally agreed."

The making of the DTCE followed a set of methodological criteria pre-established by ILTEC's working team, namely:

1. Study of the e-commerce area:
   – analysis of e-commerce vocabulary;
   – bibliographical research;
   – statement of goals and working plan;
   – evaluation of needed resources;
   – specialists' selection.

2. Preparation of the Dictionary:
   – structure of the domain's conceptual area;
   – extraction of information;
   – elaboration of the terminological records.

3. Elaboration of the Dictionary:
   – search of terms in Internet and in specialized journals;
   – search in dictionaries within the domain;
   – gathering the corpus;
   – establishment of the nomenclature;
   – filling in the terminological records;
   – consult of the specialists.

4. Linguistic pre-revision:
   – first general revision by the ILTEC's working team.

5. Specialists pre-revision:
   – first general revision by the ahp's specialists team.

6. Revision and normalization:
   – definitive revision of the terminological records by the DTCE's team (ILTEC and ahp).

7. DTCE's presentation.

Presently the DTCE is undergoing its first general revision by the ahp's specialist group. We anticipate on presenting the final version of this work before the end of the year.

The specificity of the e-commerce domain was responsible for a set of theoretical and methodological options. Referring to the establishment of the nomenclature we have considered eligible the terms found in journals, websites and glossaries of the specialty. However, the collected list of terms was submitted to a revision where we have tried to understand the integration degree of each one of the terms in the e-commerce domain. In

order to classify the collected terms we have asked for the specialists' help, arriving at the following type of terms:

– terms to reject: terms that despite occurring quite frequently in the consulted sources, do not present a specificity in the e-commerce vocabulary. These terms belong to other vocabularies. Examples: *negócio*, *transacção, empresa, consumidor.*

– terms to select: terms that occur quite frequently in the consulted sources, present specificity in the e-commerce vocabulary and were selected by specialists. It is possible to isolate these terms in two different subtypes:
a) specific terms of e-commerce. Examples: *comércio electrónico a retalho, comércio electrónico entre empresas (B2B), cibercomprador*;
b) terms belonging to several domains. Examples: *cartão inteligente, vantagem competitiva, pin de segurança, autorização em linha, serviço digital.*

In order to develop this kind of work the specialists' validation of terms is crucial, as we can see in Cabré (1999: 136-137) "not all terms occurring in the corpus need to appear in the terminology. Some terms do not belong to the topic at issue but rather to another field; in other cases, they belong to the general language. The decision whether a term is pertinent is not straightforward because the project's objectives and users are factors that can alter the choice of the units appearing in the terminology."
Having established the DTCE's nomenclature, we filled the terminological records (cf. figure 2). It was precisely in this working phase, where a considerable amount of unpredictable cases appeared, that lead to the establishment of some criteria applied to each specific case.
Concerning the insertion of the entries we have considered its reliability level, selecting it according to the available options (cf. 3.), even the terms considered to be avoided[4] were inserted referring these to a synonym considered more accurate in Portuguese.
The entries were always written with lowercase, excluding proper names, the initials were written with capitals excluding points and spaces, the subdomains were always written in capitals.

Portuguese term:	banner publicitário
Morpho-syntactic category:	S. m.
Reliability level:	term to be avoided
Synonymous:	tira publicitária
English term:	advertising banner

Table 1 – Example of a term to be avoided.

Portuguese term:	pagamento on-line
Morpho-syntactic category:	S. m.
Reliability level:	term to be avoided
Synonymous:	pagamento em linha
English term:	e-payment

Table 2 – Example of a term to be avoided.

In the next case we have two terms referring to the same English term, but both grammatical and both in use. *E-business* originated to *e-negócio* (where this *e-* is a formation element, hence it occurs at the left of the noun) and *negócio-e* (where this *-e* is a reduction of

---

[4] For "term to be avoided" we understand those that, although they can be well formed and in use in this domain, are not originally from Portuguese, as in <banner>.

electrónico). This co-occurrency of two terms denominating a same English term is the direct proof of the system's instability in the e-commerce domain, as we can see in:

Portuguese term:	e-negócio
Morpho-syntactic category:	S. m.
Reliability level:	well formed term
Synonymous:	negócio electrónico; negócio-e; e-business
English term:	e-business; electronic business

Table 3 – Example of a well formed term.

Portuguese term:	negócio-e
Morpho-syntactic category:	S. m.
Reliability level:	well formed term
Synonymous:	negócio electrónico; e-negócio; e-business
English term:	e-business; electronic business

Table 4 – Example of a well formed term.

For each term, we have tried to find a registered definition, looking for the definitions existing in the consulted sources (always identified), improving and complementing them, as we can see in the following example:

Portuguese term:	modo de pagamento electrónico
Morpho-syntactic category:	S. m.
Reliability level:	proposal for term
Synonymous:	sistema de pagamento electrónico
English term:	electronic payment mechanism
Subdomain:	Custos e modalidades de pagamento
Definition:	Modo de transferência de valores ou de pagamento de dívidas através de ferramentas electrónicas.
Definition (source):	http://www.anu.edu.au/people/Roger.Clarke/EC/EPM (16/10/2003)

Table 5 – Example of a definition.

Portuguese term:	desintermediação
Morpho-syntactic category:	S. f.
Reliability level:	well constructed term
Synonymous:	–
English term:	desintermediation
Subdomain:	Tipos de transacção
Definition:	Processo de eliminação de intermediários. No contexto da Internet, corresponde à anulação de funções ou serviços como a distribuição ou o recrutamento, uma vez que passa a haver contacto directo entre fornecedores e clientes.
Definition (source):	http://www.canalebiz.com/glossario.php?idletra=D (17/09/2003)

Table 6 – Example of a definition.

The context does not necessarily have to be definitional, but it must be chosen depending on the transparency which allows us to understand the use of the term. When a "proposal for term" has been registered in Portuguese, and the term has not yet, for this reason, been attested, there is no context.

Portuguese term:	micropagamento
Context:	"O objectivo é ampliar o uso do celular, por meio da oferta de novos serviços, especialmente os de pagamento de pequenos valores (**micropagamento**) (...)"
Context (source):	http://www.revistaecommerce.com.br/apresenta2.php?pag_id=143&edicao5=24 (16/10/2003)

Table 7 – Example of a context.

Portuguese term:	tecnologia fiável
Context:	" (...) a importância de disponibilizar **tecnologia fiável** para suportar o seu negócio na Internet".
Context (source):	www.netlojas.com/Como__/como__.html (22/09/2003)

Table 8 – Example of a context.

In identifying the websites used as sources (of definition or context), we have always indicated the consultation date given the volatility of these kind of sources.

In some cases, during the research, we have identified the use of different terms in European Portuguese (PE) and in Brazilian Portuguese (PB) as equivalents for the same term in English. In these cases the term in PE was included in the first place, giving the PB term as a synonym and then as an entry. We also found, in other occurrences, definitions and/or contexts only in PB, inserting in these cases the context with the respective indication of PB.

Portuguese term:	cartão de pagamento electrónico
Reliability level:	well formed term
Synonymous:	carteira electrônica (PB)

Table 9 – Example of a PB synonymous.

Portuguese term:	botão de adição
Reliability level:	well formed term
Synonymous:	botão de acrescentar (PB)

Table 10 – Example of a PB synonymous.

Orthographic and typographic instability is another characteristic of the e-commerce vocabulary, since the terms are allogenic units, from a structurally different linguistic system, therefore hardly being integrated in the target language.

One of the most evident cases of such instability is the "on-line" ("*em linha*" in Portuguese) case that is attested in Portuguese with the following variations: <on-line>; <online>; <"on-line">; <"online">; *<on-line>; <online>; <"on-line">; <"online">*. With this particular term we have systematically opted for the use of the form <on-line>, which is registered in one of the reference dictionaries in European Portuguese (the *Dicionário da língua portuguesa contemporânea* – DLPC). However, despite already being registered in the DLPC, we have considered that <em linha>, also registered in this dictionary, would be a more accurate term, unambiguous and frequently used in this domain of speciality. Hence, we have inserted two different terminological records each one presenting a different reliability level.

Portuguese term:	consumo em linha
Morpho-syntactic category:	S. m.
Reliability level:	well formed term
Synonymous:	consumo on-line
English term:	on-line consumption

Table 11 – Example of a well-formed term.

Portuguese term:	consumo on-line
Morpho-syntactic category:	S. m.
Reliability level:	term to be avoided
Synonymous:	consumo em linha
English term:	on-line consumption

Table 12 – Example of a term to be avoided.

One other option concerning the reliability level is the "proposal for term". Every time a term should be avoided, or yet without an equivalent in Portuguese, we have proposed a term, approved by the specialists afterwards.

Portuguese term:	sistema de controlo automatizado
Morpho-syntactic category:	S. m.
Reliability level:	proposal for term
Synonymous:	–
English term:	automated inventory control system

Table 13 – Example of a proposal for term.

Portuguese term:	custo por clique
Morpho-syntactic category:	S. m. sing.
Reliability level:	proposal for term
Synonymous:	–
English term:	cost per click

Table 14 – Example of a proposal for term.

Other example of this orthographic and typographic instability is the case of the English term "one-to-one marketing", for this term we can find five different equivalents in Portuguese such as: "marketing personalizado"; "marketing one-to-one"; "marketing um-para-um"; "marketing de um para um" and even "one-to-one marketing".

The e-commerce vocabulary has a considerable number of initials, from those representing the existent relations between the participating agents in the commercial transactions to those related to the fact that this is a commerce developed on the Internet.
As we can see in figure 3, the existent relations between companies, consumers and public administration are described through initials coming directly from English, and also used in Portuguese.
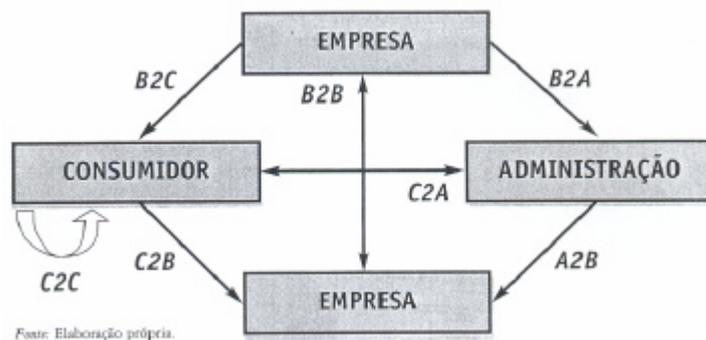

Figure 3 – Initials representing the existent kinds of relation between the e-commerce intervenient agents.

In order to represent the initials in DTCE we have inserted two entries: one representing the initial as "Portuguese term" and the deciphered initial as "synonym" (cf. table 13), and the other representing the deciphered initial as "Portuguese term" and the initial as "synonym" (cf. table 14).

Portuguese term:	B2B
Morpho-syntactic category:	S. m.
Reliability level:	term to be avoided
Synonymous:	comércio entre empresas
English term:	B2B

Table 15 – Example of a initial.

Portuguese term:	comércio entre empresas
Morpho-syntactic category:	S. m.
Reliability level:	well formed term
Synonymous:	B2B
English term:	B2B

Table 16 – Example of a deciphered initial.

As mentioned before, since the majority of e-commerce terms comes from English, we have given a particular attention to the effective use of terms in Portuguese, trying to establish an appropriate nomenclature according to the real use of terms. In the case of terms used in English we have tried to understand if and how these terms are used in Portuguese either with the help of the specialists or of the consulted written sources. Thus, we present some examples revealing that "terminography must not be confused with translation" (Cabré 1998: 115).

Portuguese term:	barreira de segurança
Morpho-syntactic category:	S. f.
Reliability level:	well formed term
Synonymous:	firewall de segurança
English term:	secure firewall

Table 17 – Example of a term without a direct translation.

Portuguese term:	directório de conteúdos
Morpho-syntactic category:	S. m.
Reliability level:	well formed term
Synonymous:	agregador de conteúdos
English term:	content aggregator

Table 18 – Example of a term without a direct translation.

5. Conclusion

With this paper we intended to present the process of elaboration of the DTCE. Thus, after the characterization of e-commerce speech, we have presented this dictionary revealing how it is structured.

Methodological questions involved in the construction of this dictionary were the aim of a particular enhance, presenting the methodological criteria established by ILTEC linguistics' team before the beginning of a work that involves the creation of a dictionary with such characteristics. We also took into account the theoretical and methodological options adopted in the filling of the terminological records in relation to the nomenclature establishment, the entries insertion, the reliability level, the used synonyms, the definitions

and contexts' selection, the differences between PE and PB, the graphical instability and the initial presentation.

With the DTCE, this team aims at contributing to the creation and/or settlement of a terminology in Portuguese based not only in technical criteria, but also in linguistics. Hence, the interaction between linguistic work and technical revision by specialists is vital to a successful result.

6. Bibliography

ANTUNES, M. & S. CORREIA (in print). «Semantic nets in the Net». *In:* Proceedings of the *XVII International Congress of Linguistics*, Prague, July 2003.

ANTUNES, M., S. CORREIA & R. GONÇALVES (2003). «E-Termos: descrição e hipótese de classificação». *In*: Proceedings of the *XVIII Encontro da Associação Portuguesa de Linguística*. Lisbon: APL, pp. 121-130.

CABRÉ, M. T. (1998). *Terminology - Theory, methods and applications*. Amsterdam / Philadelphia: John Benjamins Publishing Company.

CABRÉ, M. T. (1994). «Terminologie et dictionnaires». *In*: Meta, vol. 39, nr. 4, pp. 589-597.

CABRÈ, M. T., J. FREIXA e E. SOLÉ. (2000). *La Neologia en el Tombant de Segle*. Barcelona: Institut Universitari de Linguística Aplicada / IULA / Universitat Pompeu Fabra.

CASTELEIRO, J. Malaca. (Coord.). (2001). *Dicionário da Língua Portuguesa Contemporânea*. Lisbon: Academia das Ciências de Lisboa and Editorial Verbo.

CORREIA, M. (1998). «Neologia e Terminologia», *In*: *Terminologia: questões teóricas, métodos e projectos*. Lisbon: Publicações Europa-América, pp. 59-74.

COSTA, J. A. & A. S. Melo (1999). *Dicionário da Língua Portuguesa,* 8th. ed. Oporto: Porto Editora.

FRANQUESA I BONET, E. (dir.) (2001). *Diccionari d'Internet*. Barcelona: TERMCAT – Centre de Terminologia. Enciclopèdia Catalana.

SIEBEL, T. M. (2002). *Princípios de eBussiness*. Paris: Maxima.

TERMCAT (1990). *Metodologia del treball terminològic*. Barcelona: Generalitat de Catalunya. Departament de Cultura.