

The Phraseology of General Use ICT Texts and its Translation from English into Romanian – A Corpus-based Analysis

La phraséologie des textes à usage général du domaine de l'informatique et sa traduction de l'anglais en roumain – une analyse basée sur corpus

Frazeologia textelor de uz general din domeniul informaticii și traducerea acestora din engleză în limba română – o analiză bazată pe corpus

Sorina POSTOLEA

“Petre Andrei” University of Iași; “Al.I. Cuza” University of Iași
13, Grigore Ghica Vodă str., Iași; 11, Carol I Blvd., Iași
sorinapostolea@gmail.com

Abstract

In the Romanian area, many scholars have analysed the language specific to the field of information and communication technology (ICT) in relation to only one side of it, terminology, and to one type of terms, borrowings from English. The other features of this very rich and dynamic field of knowledge have remained quite untapped. Based on a parallel corpus of actual general use English texts and their translations into Romanian, this article aims to outline the main types of specialised collocations used in the English language of ICT as well as the main strategies used in their translation into Romanian.

Résumé

En Roumanie, beaucoup de chercheurs ont analysé le langage spécifique au domaine des technologies de l'information et des communications (ICT) en relation avec seulement une de ses dimensions, la terminologie, et avec un type de termes, les emprunts à l'anglais. Les autres caractéristiques de ce domaine du savoir très riche et dynamique n'ont pas été puisées. A l'aide d'un corpus parallèle qui contient des textes anglais à usage général et leurs traductions en Roumain, cet article se propose de relever les principaux types de collocations spécialisées employées dans l'anglais de l'informatique et les principales stratégies employées pour les traduire en roumain.

Rezumat

În spațiul românesc, numeroși cercetători au analizat limbajul specific domeniului tehnologiei informației și comunicațiilor (ICT) în relație cu doar una dintre dimensiunile sale, terminologia, și numai un tip de termeni, împrumuturile din engleză. Celelalte trăsături ale acestui domeniu al cunoașterii foarte bogat și dinamic nu au fost suficient explorate. Cu ajutorul unui corpus paralel compus din texte de uz general în limba engleză și traducerile acestora în română, acest articol își propune să evidențieze principalele tipuri de colocații specializate folosite în engleza din domeniul informatic și principalele strategii folosite pentru a le traduce în limba română.

Keywords: *specialised translation, language of ICT, terminology, phraseology, collocation*

Mots-clés: *traduction spécialisée, langage informatique, terminologie, phraséologie, collocation*

Cuvinte cheie: *traducere specializată, limbajul informatic, terminologie, frazeologie, colocație*

1. Introduction

As a language, English has always maintained a symbiotic relationship with the field of information and communication technology (ICT). Historically, the main discoveries relative to computers and their use were made in American and British universities, while the first electronic devices were successfully mass produced by American companies. With the advent of the World Wide Web, ICT has played a decisive role in the establishment of English as the new *lingua franca* used in international communication. This is why today, from a linguistic perspective, English continues to be the “language of birth” of the vast majority of terms and texts produced in the field of ICT. However, ICT is also a worldwide industry and branch of research, and, as a result, it needs to reach an extremely varied array of end-users, who spread across borders and speak many languages. It is in this process that translation comes to play a decisive role.

Romania is no exception to this rule. Like in many other countries in which English is not an official vernacular, the vast majority of Romanian users of computers and electronic gadgets must rely on translations to get information about their digital devices. It is on some of these *general use* texts produced in this process that this article focuses. More precisely, our study is based on a corpus of 62 text pairs (English source texts and their corresponding Romanian translations) that belong to three textual genres: *online news articles*, *user manuals*, and *press releases*. Our corpus was analysed with a special computer program, a *concordancer* [1] called *ParaConc*.

TOPICS: hardware products and general technologies					
TIME FRAME: September-November 2010					
TEXT GENRE	Size (in words)		Nu mber of Tex t Pairs	Target text from	ode
	N	O			
User Manuals	9.828	0.242	5	Corporate websites	M
News articles	7.178	2.739	34	www.chip.ro	HIP
	.985	.683	8	www.computer world.ro	OMP
Press Releases	.057	.340	15	Corporate websites	R
CORPUS SIZE	9.048	5.004	62		

Table 1: Corpus structure and size

Although the influence of English on the Romanian language of ICT has already been analysed by various scholars in Romania, most of these studies have focused only on the terminology of this field and only one type of terms, *borrowings*. Yet, although terminology is its most salient feature, a specialised discourse is not made up only of terms. It also encompasses a series of more or less fixed expressions, of combinations of words which are usually used together in a pre-determined manner but which are not terms per se. They are usually referred to as *phraseology*. It is on the phraseology of the general use texts included in our corpus and on the way in which it is usually translated into Romanian that this article focuses. After an examination of the

main features that define phraseology in a specialised language, our analysis puts forth a classification of the main types of phraseological units found in our corpus as well as a sketch of the key strategies used in their translation into Romanian.

2. Terms vs. collocations

Currently, almost every study concerned with phraseology [2], regardless of whether it refers to general or specialised discourse, starts with an attempt to clarify the terminological fuzziness that surrounds this concept. In fact the same reality is known under many names and several criteria are used to define it. As Anita Naciscione puts it, “it is a challenge to write on phraseology since it is an area with a confusing range of terminology and different approaches” [3]. A.P. Cowie, for instance, sums up some of the terms used to refer to this linguistic concept: *phraseological units*, *word-combinations*, *phrasal lexemes*, *phraseological expressions*, *set combinations*, *set groups*, *phrasemes*, *idioms*, *collocations*, etc. [4]. To these we may add denominations such as *multiword lexical units*, *fixed phrases*, *fixed expressions*, *prefabs*, or *phraseologisms* [5].

The boundaries and the relationships established between these concepts are not always clear-cut. For instance, from a lexical-semantic perspective and within his Meaning Text framework, Igor Mel'čuk envisages the field of phraseology as a pyramid in which the superordinate concept is that of *set phrase* or *phraseme*, defined simply as “a phrase which is not free”, meaning that, in its construction, either the selection of its constituent members or their combination, or both, are not made freely, but are restricted to a limited number of choices [6].

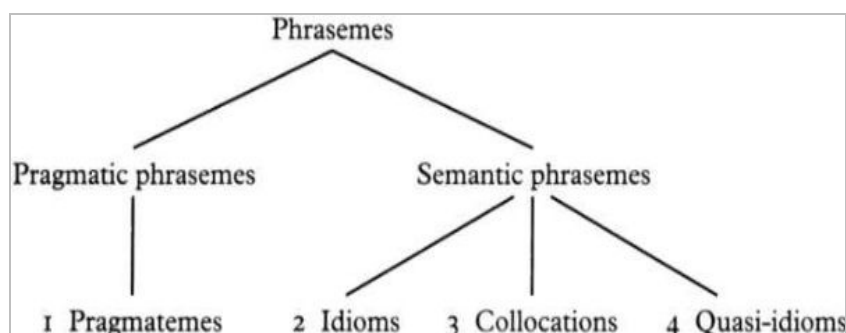


Figure 1: Typology of phraseology according to Mel'čuk (1998)

Starting from the premise that “people speak in set phrases, rather than in separate words” [7] Mel'čuk further divides phrasemes into pragmatic phrasemes, or pragmatemes, and semantic phrasemes. He includes in the former category “all ready-made expressions (like greetings, typical phrases used in letters, conversational formulae, technical clichés, proverbs, sayings, etc.)”, which “even if they are wholly compositional semantically and syntactically”, “are non-compositional pragmatically” [6]. The latter category, that of semantic phrasemes, includes *full phrasemes* or *idioms* (whose meaning is different from the sum of the signifieds of their constituent parts: *to spill the beans*), *semi-phrasemes* or *collocations* (in which the meaning of one phrase member remains intact: *to launch an attack*), and *quasi-phrasemes* or *quasi-idioms* (whose overall meaning includes the signifieds of both constituent lexemes, but also contains an unpredictable addition: *to start a family*) [9].

From a similar semantic perspective, Rosemarie Gläser (1998) puts forth a different description of phrasemes. She defines the “phraseological unit” as a “lexicalized, reproducible bilexemic or polylexemic word group in common use, which has relative syntactic and semantic stability, may be idiomatized, may carry connotations, and may have an emphatic or intensifying function in a text” [10].

More recently, adding a stylistic dimension to her analysis and drawing on the works of A.V. Kunin, Anita Naciscione (2010) defines the phraseological unit (PU) as “a stable, cohesive combination of words with a fully or partially figurative meaning”. She distinguishes between the “base form”, the “core use”, and the “instantial use” of a phraseological unit, and she claims that the two inherent properties of PUs, stability and figurative meaning (to which she adds a third – cohesion), differentiate them from “free word combinations and also from set expressions which are stable but which have no figurative meaning”. In line with Kunin, Naciscione calls the latter “stable word combinations of non-phraseological character” or “non-figurative set phrases” [11].

However, the three views described up to this point focus on the semantics of phraseology, and they use the dichotomy compositionality/non-compositionality of meaning to distinguish between the free combinations of words and the more or less fixed phrases used in a language. Yet, the definition advanced by Stefan Th. Gries introduces a new dimension into the equation, the concept of *co-occurrence*: “a phraseologism is defined as the co-occurrence of a form or a lemma of a lexical item and one or more additional linguistic elements of various kinds which functions as one semantic unit in a clause or sentence and whose frequency of co-occurrence is larger than expected on the basis of chance [12]. In fact, the definition provided by this scholar reconciles the previous views, which epitomise the traditional approach to the study of phraseology, with a relatively new trend, or, more precisely, with a “methodological paradigm” [13], that of corpus linguistics.

Thus, one of the most prominent scholars in corpus studies and lexical semantics, Michael Stubbs, uses *phrase* as a “neutral term for a string of words” which constitutes an “extended lexical unit”, that is, a unit of meaning. In his acceptance, a collocation is “a lexical relation between two or more words which have a tendency to co-occur within a few words of each other in running text” [14]. Further on, he refines his view when he states that, in fact, the concept of collocation refers to the “the co-occurrence of words”: “usually it is frequent co-occurrences which are of interest, and corpus linguistics is based on the assumption that events which are frequent are significant. My definition is therefore a statistical one: 'collocation' is frequent co-occurrence” [15]. In Stubbs’s system a collocation consists of a node-word (a word-form or a lemma [16]) and of one or several collocates (word-forms or lemmas) distributed to the left and/or right of the node in a string of text.

Finally, a definition which encompasses several aspects discussed above and supplements them with a morphological-syntactic dimension is that provided by Willy Martin: “A collocation is a word group consisting of two conceptual elements: a collocator (determinans) and a base (determinatum), of a certain syntactic type (N+N, V+N, A+N, Adv+V, Adv+A), showing a semantic, type-based, relationship between the two elements; the more the collocator is, conceptually speaking, type-bound (bound to a conceptual type or category) and the more it is, lexically speaking, token-bound (bound to a lexical token/item), the more we are dealing with a collocation that conceptually and lexically forms a unit, i.e. one that is a lexical collocation [17].

However, if this definition seems to be suited for general discourse, how could one define collocations in a specialised field of knowledge and activity? A distinguished terminologist, M.T. Cabré describes the phrasal entities particular to a specialised language as a “series of constructions that are neither phrasal terms proper nor totally free syntactic formations”. They are “combinations that occur frequently in the discourse of a special subject field (which makes them similar to terminological phrases), but they do not correspond to established concepts” [18]. However, her definition seems to be rather vague, since it could encompass not only terminological phrasemes, but also all other sorts of collocations: among the combinations that occur frequently in the discourse of a special subject field and which do not refer to concepts one could find phrases from general discourse or from other specialties (in an academic paper on a specialist topic, for instance, there could be rhetoric formulae, in an ICT user manual, phrasemes relative to legal liabilities, etc.)

Yet, for the purposes of our study, a more comprehensive frame of definition is provided by Guadalupe Aguado de Cea, who establishes a more unambiguous relationship between

terminological units and phrasemes. In her opinion, specialised phraseological units (SPUs) are mainly defined by seven features: a) they are syntagmatic structures that include a term or a terminological unit; b) they include a verb or a deverbal element; c) they have a high degree of fixedness, although sometimes they allow the insertion of other elements, such as a quantifier or a pronoun; d) they demonstrate regularity in their composition; e) they admit the commutation of their constituents, but to a certain degree; f) they maintain a specific meaning within a given field, and thus they acquire semantic stability; g) they are of frequent use in a specialised field of subfield of activity [19].

Based on the notions discussed in this section, in what follows we present an overview of the main types of collocations discernable in our specialised ICT parallel corpus.

3. Typology of collocations in general use ICT texts

The SPUs specific to the field of information and communication technology may be analysed from various points of view, such as their meaning, their origin, their morphology, or a combination of these, etc. However, two criteria seem to be more relevant in this context: the lexical/terminological one, and the morphological/syntactic one.

According to the former criterion, a comprehensive general typology of phraseological units, which encompasses both the lexical-semantic and the corpus-based view presented in the previous section, was provided by Sylviane Granger and Magali Paquot [20]. According to the functions they perform in discourse, they divide the vast category of phrasemes into referential, textual and communicative.

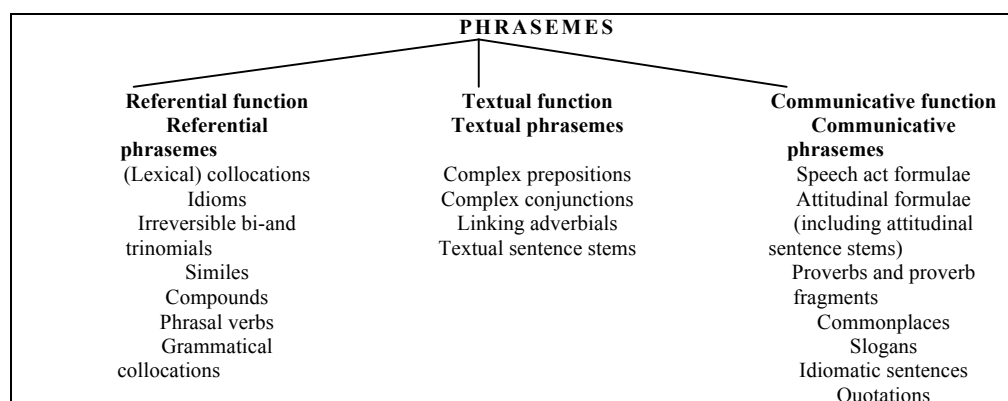


Figure 2 : The phraseological spectrum according to Granger & Paquot (2008)

As we have embraced Aguado de Cea's opinion according to which a SPU or specialised collocation necessarily includes a term, it is clear that our typology of the main phraseological units present in our corpus should focus on the first category of phrasemes described in *Figure 2*, namely, *referential phrases*. Although present, to a given extent, in our texts, the other types of phraseological units (textual and communicative) are not within the scope of this study because they are not field-bound but they rather belong to textual conventions and should be analysed by text-type, from a stylistic and functional perspective.

Subsequently, from a lexical perspective, the SPUs in our ICT corpus could be divided into:

a) **LEXICAL COLLOCATIONS** – “usage-determined or preferred syntagmatic relations between two lexemes in a specific syntactic pattern. Both lexemes make an isolable semantic contribution to the word combination but they do not have the same status” [21]. This definition of collocations draws on that put forth by Mel'čuk and it would not be farfetched to claim that it is probably the most frequent type of SPU found in a specialised field. In our corpus, examples of terminological collocations include: *to start a program*; *Internet browsing*; *page layout*; *system*

settings; system failure; to setup a program; to rip a file; data processing, text editing, memory run; etc.

b) IRREVERSIBLE BI-AND TRINOMIALS – “Irreversible bi- and trinomials are fixed sequences of two or three word forms that belong to the same part-of-speech category and are linked by the conjunction ‘and’ or ‘or’” [22]. In our corpus, examples of such units in the language of ICT include: *hardware and software; left and right* (buttons, hand sides, speakers); *drivers and utilities; clear and crisp*, etc.

c) COMPOUNDS – “Compounds are morphologically made up of two elements which have independent status outside these word combinations. (...) They resemble single words in that they carry meaning as a whole and are characterized by high degrees of inflexibility, viz. set order and non-interruptibility of their parts” (*ibidem*). In the field of ICT, such compounds include: *broad-band; built-in; hand-held; Internet-connected; high-tech*; etc.

d) PHRASAL VERBS – as some authors have already noticed, a series of verbs which have acquired a stable specialised meaning in a specialist field may be called terminological verbs; in this case, such ICT terminological phrasal verbs are: *to boot up* (a computer); *to back up* (data); *to set up* (a system); *to sign in/off*;

Idioms and similes were not included into this typology because they belong to other uses of language and are absent or exceptionally scarce in specialised discourse. Grammatical collocations do not fall within the scope of this study.

Another way to look at specialised phraseological units is from the angle of their morphological/syntactic structure. Based on Aguado de Cea’s [23] classification of SPUs, we may divide the terminological collocations in our corpus into the following structures:

a) V + N (or prepositional phrase): *navigate the menu; start an installation; write/read data; recover data; transmit data files or messages; store in memory*; etc.

b) N + N: *barcode reading; battery charging; battery pack; data stream; data protection; read speed; word processing*;

c) Adj + N: *advanced configuration; high-tech industry; remote control*. Actually, in the language of ICT, a very prolific subtype of this phrasal structure is the following: [N+Adj/Adv] + N: *energy-efficient processors; machine readable form; user-friendly interface; battery-friendly technology; Internet-connected devices; Ethernet-based network; web-based content*; etc. Also, adjectives derived from verbs with the suffix *-able/ible* are widely used in ICT texts: *portable Wi-Fi hotspots; more portable storage and speed; free downloadable data recovery software; consumable item; non-bootable (non-system) disk; convertible PC experience*; etc.

d) N + prep. + N: *line of code; mode of operation; frames per second; performance per watt; bits per colour*; etc.

e) V + Adv – actually, in our corpus of general use ICT texts, much more frequent is the intensifying combination Adv + V: *to rapidly take successive shots; to wirelessly pair up; to seamlessly move between devices*; etc.

f) Adv + Adj: *fast processing system; highly rated applications; highly portable and durable device; increasingly capable; environmentally-friendly technology*;

As some of these combinations are foreign to the structure of Romanian as a language and they sometimes include terms that are neologisms in Romania, in the following lines we attempt to highlight the main strategies used in their translation.

4. Strategies used in the translation of ICT phraseological units

One could argue that in an ICT text, just like in the case of other types of discourse, it is phraseology which is the most difficult to translate. While there are numerous dictionaries and thesauri which offer lengthy explanations and accurate translations of terms, the phraseology of a specialist field is sometimes difficult to grasp. As Guadalupe Aguado de Cea notices, it is not terms,

but phrasemes which provide “naturalness” to scientific discourse [24]. Or, in other words, “it is often the collocation of two or more lexical items, rather than the occurrence of isolated terms that determines the identity of a given register” [25].

Before we attempt to highlight the strategies which are usually used by Romanian translators when they transfer English ICT texts, let us have a look at some collocational ranges [26] found in our corpus and at their translations:

English - ST	Romanian - TT
Base / node-term: MEMORY	
... to 1GB of <i>GDDR5</i> memory running at 1050MHz	... catre memoria <i>GDDR5</i> de 1GB ce ruleaza la 1050MHz
... optimize airflow around the modules helping <i>memory run</i> at peak performance...	... optimizeaza fluxul de aer din preajma modulelor, ajutand <i>memoria sa functioneze</i> la nivelul cel mai ridicat al performantei...
You can <i>physically expand the memory</i> up to 8 GB...	<i>Memoria poate fi extinsa fizic</i> pana la 8 GB...
The <i>memory cannot be increased</i> .	<i>Memoria nu poate fi marita</i> .
... the total amount of <i>usable memory</i> available will be less...	... <i>memoria totala disponibila utilizabila</i> va fi mai mica...
Many of the advances in the chip come from its <i>use of memory</i> .	Multe din avantajele din acest cip vor veni din <i>folosirea memoriei</i> .
... the stop position will be <i>stored in memory</i> va fi <i>memorata</i> ultima pozitie de redare.
... a large 1536MB <i>GDDR5</i> video memory, rated at over 4GHz...	... de 1536 MB <i>memorie video GDDR5</i> , tactata la 4 GHz.
<i>Memory clock</i> 4008MHz	<i>Frecventa memoriei</i> 4008 MHz
Base / node-term: FILE	
USB 3.0 speeds up <i>file transfers</i> ten-fold...	Prin interfata USB 3.0, <i>transferul de fisiere</i> va fi de 10 ori mai rapid...
... not all of the <i>recorded files may be played</i> este posibil sa nu <i>poata fi redade toate fisierele inregistrate</i> .
... only the <i>JPEG files</i> in the picture folder can be played.	...puteti <i>vizualiza doar fisierele JPEG</i> din folderul cu poze.
<i>Viewing the file</i> requires Adobe Reader...	<i>Vizualizarea fisierului</i> necesita Adobe Reader...
... a <i>DivX file</i> created by the user <i>might not be played</i> este posibil <i>sa nu puteti reda un fisier DivX creat de dvs.</i>
To play a <i>file icon</i> in the screen...	Pentru a reda <i>pictograma unui fisier</i> de pe ecran...
For <i>files encoded in Variable Bit Rate (VBR) format</i> ...	Pentru <i>fisierele codate intr-un format</i> cu rata de biti variabila (VBR)...
To <i>select all files</i> select the [] Icon...	Pentru a <i>selecta toate fisierele</i> , selectati pictograma []...
DivX files have <i>.Avi file extensions</i> , however, not all <i>.Avi files</i> are DivX...	Desi fisierele DivX au <i>extensii .Avi</i> , nu toate <i>fisierele cu extensia .Avi</i> sunt in <i>formatul DivX</i> ...
It is important to <i>back up all data files</i> before using this option.	Inainte de a folosi aceasta optiune, este important <i>sa faceti o copie de siguranta tuturor fisierelor de date</i> .

Table 2: Collocational ranges of *memory* and *file*

This is but a trivial sample of the vast collocational web encompassed in ICT texts and we should also mention that this sample deliberately excludes phrasal terms, i.e. complex terms such as *memory card*, *memory card slot*, *data file*, *video file*, etc. However, this small sample is enough to reveal some strategies used in the translation of ICT phraseological units. As the translation of phraseology develops at the sentence level, the framework of analysis provided by the linguist turn

in translation studies seems to be the best choice in order to account of the strategies used in their interlingual transfer.

Lucía Molina and Amparo Hurtado Albir [27] provide a revised functional classification of traditional translation techniques [28]. According to this framework, a provisional classification of the main strategies and techniques used in the English-Romanian translation of ICT collocations would include the following items:

a) BORROWING. This is a widely used technique, both at the level of terms and at that of phrases. From the viewpoint of their degree of naturalisation, phraseological borrowings may be divided into:

- full: *Select Driver Setup* → *Alegeti Driver Setup*;

wireless keyboard and mouse → *tastatura sau mouse wireless*;

- integrated: *Double click time* → *Intervalul dublu click*;

turn off all wireless or radio transmitting devices → *opriti toate dispozitivele de transmisie radio sau wireless*;

boot screen → *ecranul de boot-are*;

- naturalised: *Mouse settings* → *setarile pentru maus*;

with one click → *printr-un singur clic*

b) CALQUE. At this level of analysis it is sometimes difficult to distinguish between calques and the next technique in our typology, literal translation. Thus, we make a distinction between calque, used in case of bi-lexical phrasal units, and literal translation, used in the case of larger word combinations. Examples of calques include:

available for download → *disponibil spre descarcare*

recorded files → *fisiere inregistrate*

data encryption → *criptarea datelor*

to navigate the menu → *sa navighezi in meniu*

touch-sensitive buttons → *butoane tactile-senzitive*

c) LITERAL TRANSLATION. This technique applies to larger ICT phraseological units, which are translated with minimum changes that reflect only the specific patterns of Romanian as a Romance language.

transmission of data files or messages → *transmisia de fisiere de date sau mesaje*

a highly portable and durable device → *un dispozitiv foarte portabil si durabil*

d) AMPLIFICATION. In this procedure, the translator makes explicit information which is implicit in the TT:

data recovery software download from → *aplicatie de recuperare a datelor, disponibilă pentru download de la*;

clicking the right mouse button → *faceti clic pe butonul aplicatiei dreapta al mausului*

e) REDUCTION. In this procedure, the translator omits some information which is explicit in the source text.

operators have to carry the data stream → *operatorii transmit datele*

Accessory Bundled → *Accesorii*

battery-friendly NVIDIA Optimus technology → *tehnologia NVIDIA Optimus*

f) LINGUISTIC AMPLIFICATION. Linguistic amplification occurs when the target segment of language encompasses more elements than in the source text. Leaving aside the numerous situations in which the added elements are “required” by the structure of the target language (for instance, adding the preposition *de* when translating English N + N structures, as in *data file* → *fisier de date*), examples of this technique include:

click (verb) → *faceti clic*

It is important *to back up* all data files → *este important sa faceti o copie de siguranta tuturor fisierelor*

g) LINGUISTIC COMPRESSION. Basically, when translators use this procedure, they say the same thing with a smaller number of words.

stored in memory → *memorată*

Rated power consumption → *Consum nominal*

h) ESTABLISHED TRANSLATION. Although these cases are quite rare, the Romanian language of ICT includes some terms and collocations which have established translations that are neither calques, nor literal translations of their source counterparts. Examples in this respect are the terms *clock* [29] and *zip*, and their compounds and collocates.

the 1015PEM *has clocked* at speeds... → 1015PEM *este tactat* la viteze...

factory overclocked → *supratactata din fabrica*

built-in unzip utility → *programul de dezarhivare integrat*

i) GENERALISATION. This technique involves using a more general term to express a particular item of the source text.

battery pack → *baterie*

revolutionary new *sunlight screen technology* → *ecran* revoluționar

with a *screen size* of → *cu ecran* de

j) PARTICULARISATION. Of course, this procedure is the opposite of generalization and it involves using a particular term to express a general item of the source text.

battery pack → *acumulator*

portable electronic devices → *gadgeturi portabile*

k) MODULATION. In this procedure, the point of view originally expressed in the source text is modified. A good example in this respect is the following:

High power efficiency means *greener, more environmentally-friendly technology* →

Eficiența ridicată înseamnă *protejarea mediului inconjurator*

can process a floating point operation → *poate procesa în virgulă mobilă*

l) TRANSPOSITION. This technique involves changing the grammatical category or categories of the source phrasal unit:

like a jukebox, with *on demand browsing* → *ca un tonomat, care selectează la cerere*

m) DESCRIPTION. In this technique, the target unit is a description or an explanation of the source segment.

twin, long-lasting, *hot-swappable batteries* → *bateriile sale duble, durabile, care pot fi schimbate în timpul funcționării*

including support for *four-finger multiple touch* → *acceptând inclusiv atingerea a patru degete în același timp*

Of course, examples of 100% pure translation techniques are quite difficult to find, and the boundaries between them are sometimes hard to establish (for instance, between description and simple/linguistic amplification). In the vast majority of cases, the translation of each source-text phrasal unit involves several techniques and decisions from the part of the translator. It goes without saying that these strategies depend on the nature of the source text as a whole, on stylistic genre-based criteria, on linguistic and terminological restrictions, and on various factors related to the process of translation itself (such as the deadline, the conditions in which the translator works, etc.). For instance, the following Romanian translation uses at least three procedures which involve specialised phraseological units, as shown below:

DivX files have .Avi file extensions, however, not all .Avi files are DivX...

Desi fisierele DivX au extensii .Avi, nu toate fisierele cu extensia .Avi sunt in formatul DivX...

Calque / borrowing: *DivX files* → *fisierele DivX*

Linguistic compression / generalisation: *.Avi file extensions* → *extensii .Avi*

Amplification: *.Avi files* → *fisierele cu extensia .Avi or DivX* → *in formatul DivX*

5. Conclusions

Although it used a only small-size corpus of general use ICT texts, and, because of the limited space available, it focused on only a handful of examples, our analysis has shown, in our opinion, both the importance of phraseology in the study of specialised languages and its extremely rich potential in translation-related research. Even though terms and terminology usually get the “lion’s share” of the researchers’ interest, collocations and phraseology seem to be just as important since, in the actual practice of translation, they could become a source of difficulty for translators and they are often a marker of a translation’s “naturalness” in the target language.

References

- [1] In corpus linguistics, computer programs purposefully designed for the linguistic analysis of corpora are called *concordancers* or *concordance programs*
- [2] This term refers both to the branch of linguistics concerned with the analysis of these parts of language and its object of study.
- [3] NACISCIONE, Anita, *Stylistic Use of Phraseological Units in Discourse*, John Benjamins Publishing Company, Amsterdam / Philadelphia, 2010, p. 17.
- [4] COWIE, A. P. “Introduction” in *Phraseology. Theory, Analysis, and Applications*, Clarendon Press, Oxford, 1998, *passim*.
- [5] NACISCIONE, Anita, *op.cit.*, p. 17.
- [6] MELČUK, Igor, “Collocations and Lexical Functions”, in A. P. Cowie [ed.], *Phraseology. Theory, Analysis, and Applications*, Clarendon Press, Oxford, 1998, p. 27.
- [7] *Ibidem*, p. 24.
- [8] *Ibidem*, p. 27.; In the literature and from a semantic perspective, the criterion of non-compositionality is often used to discriminate between phrasal and non-phrasal chunks of language. Lyons [30] explains compositionality by stating that “the meaning of a composite expression is a function of the meanings of its component expressions”. In layman terms, the meaning of a phraseme should differ from the “sum” of the meanings of its constituent parts: *green policy*.
- [9] *Ibidem*, pp. 29-30.
- [10] GLÄSER, Rosemarie, “The Stylistic Potential of Phraseological Units in the Light of Genre Analysis”, in A. P. Cowie [ed.], *Phraseology. Theory, Analysis, and Applications*. Clarendon Press, Oxford, 1998, p. 124.
- [11] NACISCIONE, Anita, *op.cit.*, p. 19.
- [12] GRIES, Stefan Th., “Phraseology and linguistic theory. A brief survey”, in Sylviane Granger and Fanny Meunier [eds.], *Phraseology: an interdisciplinary perspective*, John Benjamins, Amsterdam / Philadelphia, 2008, p. 6.
- [13] GLÄSER, Rosemarie, *op.cit.*, p. 15.
- [14] STUBBS, Michael, *Words and Phrases. Corpus Studies of Lexical Semantics*, Blackwell Publishing, Oxford/Malden, 2002, p. 24.
- [15] *Ibidem*, p. 29.
- [16] “Word-forms are directly observable units; a text consists of a sequence of word-forms. (...) Lemmas are not directly observable but abstract classes of word-forms; a vocabulary is usually represented as a list of lemmas” [31].
- [17] MARTIN, Willy, “A unified approach to semantic frames and collocational patterns”, in Sylviane Granger and Fanny Meunier [eds.], *Phraseology: an interdisciplinary perspective*. John Benjamins, Amsterdam / Philadelphia, 2008, p. 56.
- [18] CABRÉ, Maria Teresa, *Terminology. Theory, Methods and Applications*, in Juan C. Sager [ed.], Janet Ann DeCesaris, [trans.], John Benjamins, Amsterdam / Philadelphia, 1999, p. 91.

- [19] AGUADO DE CEA, Guadalupe, “La fraseología en las lenguas de especialidad”, in Enrique Alcaraz-Varó, J. M. Martínez and F. Y. Ramos [eds.], *Las lenguas profesionales y académicas*, Editorial Ariel, Barcelona, 2007, pp. 57-58.
- [20] GRANGER, Sylviane, PAQUOT, Magali, “Disentangling the phraseological web”, in Sylviane Granger and Fanny Meunier [eds.], *Phraseology: an interdisciplinary perspective*. John Benjamins, Amsterdam / Philadelphia, 2008.
- [21] *Ibidem*, p. 43.
- [22] *Ibidem*.
- [23] AGUADO DE CEA, Guadalupe, *op.cit.*, p. 59.
- [24] *Ibidem*, p. 53.
- [25] TROSBORG, Anna, “Text Typology: Register, Genre and Text Type”, in Anna Trosborg [ed.], *Text Typology and Translation*. John Benjamins, Amsterdam / Philadelphia, 1997 p. 6.
- [26] According to Mona Baker, a collocational range “refers to the set of collocates, that is other words, which are typically associated with the word in question” [32].
- [27] MOLINA, Lucía, HURTADO ALBIR, Amparo, “Translation Techniques Revisited: A Dynamic and Functionalist Approach”, in *Meta: journal des traducteurs / Meta: Translators' Journal*, 4, vol. 47, 2002, pp. 498-512.
- [28] These scholars distinguish between translation strategies and techniques. The former “are the procedures (conscious or unconscious, verbal or nonverbal) used by the translator to solve problems that emerge when carrying out the translation process with a particular objective in mind” while a technique is the practical and actual solution used to solve a translation problem [33].
- [29] In the field of ICT, a *clock* is “a circuit that generates pulses used to synchronise equipment” while the *clock rate* is “the number of pulses that a clock generates every second” [34].
- [30] cited in SVENSSON, Maria Helena, “A very complex criterion of fixedness: Non-compositionality”, in Sylviane Granger and Fanny Meunier [ed.], *Phraseology: an interdisciplinary perspective*, John Benjamins, Amsterdam / Philadelphia, 2008.
- [31] STUBBS, Michael, *op.cit.*, p. 26.
- [32] BAKER, Mona, *In Other Words. A coursebook on translation*, Routledge, London and New York, 1992, p. 49
- [33] MOLINA, Lucía, HURTADO ALBIR, Amparo, *op.cit.*, p. 508.
- [34] *Dictionary of ICT*, 4th edition, Bloomsbury, London, 2004.

