Yolanda-Mirela CATELLY
"Politehnica" University of Bucharest, Romania

CREATIVITY IN GENERATING COMPUTER SCIENCE TERMINOLOGY – AN ETYMOLOGICAL STUDY FROM THE APPLIED LINGUISTICS PERSPECTIVE

Motto: *'The most important thing in the programming language is the name. A language will not succeed without a good name. I have recently invented a very good name and now I am looking for a suitable language.'* (Donald Knuth - *American computer scientist*)

1. Theoretical background and main aims

The present study has been initiated starting from the observation that the creation of new words in English in the field of computer science (CS) /information technology (IT) has for some decades now been the realm of creativity in a dynamic interesting and sometimes unexpected process.

The etymological investigation in this technical domain can be both challenging and rewarding from the perspective of applied linguistics, as we will try to demonstrate in what follows. The findings from the etymological research can be turned into valuable instructional input in an ESP course for engineering students specializing in IT, from various points of view.

As pointed out in the literature (Howe, 1985), computer science terminology has been developing immensely over the last 30 years or so, resulting in the publication of newly-coined terms dictionaries amounting to over 15,000 definitions, which actually means several megabytes of explanatory text. Moreover, the appearance of these new terms is taking place in real time, while we are studying them in the English class or in the technical ones! Therefore, it is our aim to analyze them, together with our students, in search for the main patterns of creativity and imagination development underlying their formation.

We believe that for our trainees it could really be beneficial to become aware of a certain type of attitude in their professional and personal development, marked by an orientation focused on independence and creativity in thought and action. Furthermore, once the etymology of IT terminology is introduced among the points of interest of an ESP course for IT engineering students - as is the case of the research presented here, beside the advantage of sensitizing the learners as regards the linguistic and imaginative issues, there is one more beneficial aspect to be taken into consideration by the ESP teacher/course designer in CS, viz. the fact that we are dealing with a field that keeps rejuvenating, with numerous terms becoming obsolete or uninteresting quite rapidly, from the scientific/technical perspective, while at the same time numerous new others keep being created… and denominated! Consequently, the etymology of IT terminology and the instructional linguistic tasks created in this respect for class use have the advantage

of remaining valid for the students from the temporal perspective; thus, they can be successfully included in the course and be preserved in it for a much longer period of time, without the risk of losing the trainees' interest in them from their content validity point of view.

It is of significance to note at this point that, in general, etymology theories are in full agreement in accepting a range of means underlying words origins in any language, English included – with *borrowing*, various types of *word formation*, as for instance derivation and creating compounds, or *onomatopoeia* and other sound symbols among the classical ones.

Of course, challenging traditional etymological views is by no means our focus here, but rather complementing, and hopefully enriching, the existing background with an illustration of a fresh perspective on the etymology of a contemporary field of human endeavor, namely IT, for which various theories and testimonials concerning the origin of specific terms are still in wait for the researcher's attention, particularly in English, the world's recognized *lingua franca* of computing. Indeed, according to a recent study (W3Techs, 2013) concerning the presence of English in computing, in percentage of frequency, over half of the most visited websites use English as their content language (55%). This is not surprising, if we take into consideration the fact that, historically, due to technical issues in the design of the first computer generations, as well as to a massive lack of standardization on the Internet, computer professionals had to limit themselves to the use of English and the Latin alphabet. Thus, a widely quoted figure for the amount of web content in English reaches an impressive 80%. Hence, there are several elements to be taken into consideration, viz. in a non-prioritized order: a necessary awareness of the social and cultural background of those who were/are the creators of the newly coined IT terms as (non)native speakers of English, frequently living in English–speaking countries, as a factor of major influence in denominating IT specific words - at individual level, but also at organizational and societal levels. Creativity and imagination in generating IT terminology over the last decades are therefore culture-bound, and neglecting these aspects would be an error, most particularly if the findings in etymology are transferred into ESP course content, therefore they are analyzed from an applied linguistics perspective, as this is the case here.

The most recent online dictionaries (Merriam-Webster, 2014) analyzing the etymological phenomena in English for our immediate contemporary times, which necessarily includes CS/IT, list a range of categories of word formation mechanisms, among which the main processes seem to be: *borrowing, shortening/clipping, functional shift, back-formation, blends, acronymic formations, transfer of personal/place names, imitation of sounds, folk etymology, combining word elements, literary/creative coinages*.

Some comments are mandatory in this respect:

-    as can be noted, the traditional etymological means of *borrowing*, *word formation* and *sound imitation* are still to be found in our times;

-    to them, a range of newer ones have been added, out of which, according to the research of frequency carried out by this study author and her IT students/would-be specialists in the field, the ones most frequently found in the etymology of computer science terminology are: *clippings, acronyms, creative*

coinages – which speaks in itself for the creative-imaginative component in denominating IT terms, as well as an inclination towards simplification and brevity, not without an obvious touch of humor in very many cases, as most of the etymological 'stories' we have found can confirm;

-        imagination and/or humor are also to be 'guilty' of the quite numerous alleged explanations of certain acronymic etymologies, although reliable sources of citation cannot always be found for them;

-        there have been some spontaneously created words resulting from the sheer imagination of their author; however, unless such terms gain popularity and become as widely-spread as to be an entry in a dictionary, or unless the creators/their organizations are well-known or if the technical component in them is not generalized, we believe that they do not really have good chances of entering the current lexis of the field.

If one discusses the etymology of the terminology of computer science in particular – and in fact there are quite few such studies and resources, then one can note that the types of etymological explanations are restricted (The Etymology of the Terminology of Computer Science, 2014) to just two categories:

(i)   some terms are simply explained from the strict technical perspective, e.g. ''*cookie - a packet of information that travels between a browser and the web server'*, although for this term some other sources provide an extended explanation relying on metaphor and creativity, i.e. connected with the famous '*fortune cookies'* from the everyday lay English life;

(ii) some, but the proportion in the dictionary is not really in their favor, do have a story attached, which makes them belong to the creative etymological mechanism, such as the following: '*GNU - a project with a goal of creating a free operating system. Gnu is also a species of African antelope'*. To this initial explanation, two more are added to the same entry – we will give them below as an illustration of the combined character of IT terminology currently existing etymological accounts, which are imaginative, humorous and taking satisfaction from fancy acronymic explanatory narratives: (a) '*Founder of the GNU project Richard Stallman liked the name because of the humour associated with its pronunciation and was also influenced by the song The Gnu Song, by Flanders and Swann which is a song sung by a gnu'* and (b) '*also fitted into the recursive acronym culture with GNU's Not Unix.'*

We agree with the remarks on the so-called *Informatics jargon* (Topală, 2009), which mainly relies on two linguistic aspects – *metaphor* and *colloquialisms*, with emphasis being placed on a lack of standardization and uniformization, understood as a result of linguistic innovation and also of an effort to adapt to the new virtual communication situations. The author speaks even of a '*certain linguistic extravagance'*, that seems to be specific to the quite often young members of the IT community of (non)professionals.

2.   Current etymology mechanisms in IT terminology - a synthetic approach

In what follows, an approach is attempted at, which is meant to illustrate the various etymological categories to be found, in our opinion, in the current IT terminology. Needless to remark from the very beginning that it has been really difficult to strictly delineate the category/-ies of etymological means conducive to

the creation of each IT terms discussed below. On the contrary, we have tried to emphasize the existing combinations of components, as evidence of: the (omni) presence of *humor*, need for *brevity* and *creativity* in coining most of them.

As the domain is being analyzed almost in real time, it is important to mention that one can find different explanations for the same term, in function of:

(i) the type of source – technical or popular, and

(ii) of an inherent amount of subjectivity, which in most of the cases makes it difficult to classify the entries into one etymological category only.

It may be interesting to note that most of the examples come from the data obtained by means of tasks focused on thematic Internet search project work assigned to the students, while the classifications proposed are also the result of class discussions within the same approach. Table 1 below synthetically presents: a selection of typical CS/IT terms whose etymology is analyzed, the (abridged, out of lack-of-space reasons) explanation & its accredited source in each case, as well as the type(s) of etymological explanation it embodies.

Table 1. *Examples of creative etymology types identified in IT terminology*

No	Term	Explanation	Type(s) / Comments
1	ID10T	*a code frequently used by a customer service representative (CSR) to annotate their notes and identify the source of a problem as the person who is reporting the problem rather than the system being blamed. This is a thinly veiled reference to the CSR's opinion that the person reporting the problem is an IDIOT* (http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	partially acronymic & humor
2	ICQ	*an instant messaging service. ICQ is not an acronym. It is a play on the phrase "I seek you"* (http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	acronymic / pronunciation & humorous creative
3	Yahoo	*an acronym for "Yet Another Hierarchical Officious Oracle" … originally invented by Jonathan Swift and used in his book 'Gulliver's Travels'. It represents a person who is repulsive in appearance and action and is barely human* ((http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	acronymic & literary coinage
4	Apple	*after some time spent working on apple orchards in that state, he Jobs – our note) was in the midst of a* "fruitarian diet" *and thought the simple name was* "fun, spirited, and not intimidating", *the rest, of course, is history…Apple's original logo showed famous physicist Isaac Newton sitting under an Apple tree. This is an obvious reference to Newton's theory of gravitation, but now in light of the companies naming history, seems to give that apple tree a dual meaning* (http://osxdaily.com/2011/10/21/why-apple-is-called-apple)	literary coinage, (possibly speculative, as there are several versions of the name origin explanation)
5	Trojan horse	*a malicious program that is disguised as legitimate software… derived from the classical myth of the Trojan Horse* (http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	literary coinage
6	Mozilla	*When Marc Andreessen, founder of Netscape, created a browser to replace the Mosaic browser, it was internally named Mozilla (Mosaic-Killer, Godzilla)* (http://mastermason.com/WallerNews/WN0310.htm)	cultural coinage from film industry

7	Python	*from Monty Python's flying circus, a BBC comedy series from the seventies* (http://effbot.org/pyfaq/why-is-it-called-python.htm)	cultural coinage - film industry
8	Ubuntu	*a Nguni Bantu term … an idea from the Southern African region which means literally "human-ness," and is often translated as "humanity towards others," but is often used in a more philosophical sense to mean "the belief in a universal bond of sharing that connects all humanity"* (http://www.mediander.com/connects/285313/ubuntu-%28philosophy%29/#!/)	socio-cultural coinage
9	Wiki pedia	*a hawaiian word for "fast" … "wiki" can be expanded as "what I know is"* (http://foswiki.org/system/wikiculture)	borrowing, cultural coinage and also acronymic
10	Sosumi	*Apple Computer had a long litigation history with Apple Records... Fearing that the ability to record musical sound would cause yet more legal action, the Apple legal department allegedly ordered the sound to be renamed from its original, musical name. So the developers changed the name to Sosumi ("So sue me"). Depending on who was asked, they quipped that it was Japanese for either "absence of sound" or "a light pleasing tone"* (http://web.presby.edu/~phmeeker/classes/pc/CSC450/Notes/The%20Etymology%20of%20the%20Terminology%20of%20Computer%20Science.doc)	sound symbolism & humor (omophonic pun)
11	<u>bug</u>	*credited to Grace Hopper. In 1946, she joined the Harvard Faculty at the Computation Laboratory where she traced an error in the Harvard Mark II to a moth trapped in a relay. This bug was carefully removed and taped to the log book. However, use of the word 'bug' to describe defects in mechanical systems dates back to at least the 1870s, perhaps especially in Scotland. Thomas Edison, for one, used the term in his notebooks and letters* (http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	folk etymology (real case?!) & metaphor, but here imagination wins…
12	Facebook	*a colloquialism for the directory given to students at some American universities… a printed or online directory found at American universities consisting of individuals' photographs and names... it denotes publications of this type distributed by university administrations at the start of the academic year with the intention of helping students get to know each other* (http://www.askives.com/facebook-wikipedia-link.html)	extending meaning from a colloquial common noun, with a touch of humor
13	Ethernet	*an endocentric compound of "luminiferous ether" – the "substance" that was widely believed to be the medium through which electromagnetic radiation propagated in the late 19th century – and "net", short for "network"* (http://en.wikipedia.org/wiki/List_of_computer_term_etymologies)	combined word elements
14	Blu-ray	*spelt Blu instead of Blue because by law daily-used name cannot be registered as a trademark* (http://laserpointerforums.com/f38/why-blu-ray-called-blu-ray-56877.html#!/exjun)	a modified blend for legal reasons

The items presented above cannot possibly stand for all the categories and combinations of categories one can identify in IT terminology etymology, but it is not exhaustiveness we are aiming at. However, they can be illustrative as depicting the main new tendencies in the mechanisms of denominating IT words.

As can be seen from an analysis of the types included in the table, the range of categories comprises various *acronymic* forms, generally imbued with a touch of

humor, combined with *literary/cultural* and even *social* coinage, including here *film* hints, as we are in a 21st century domain of activity. *Colloquial* and *folk etymology* are also present, fighting for a top place, as in the case of Facebook, while the reality of an American English speaking culture which creates and imposes such terms also has to take into consideration *legal reasons* in coining new IT terms, as well as *social and political* concepts belonging to a society attempting to observe the ideal of transition from apartheid to democracy, in an ever more politically correct world.

3. An applied linguistics perspective on IT terminology etymology

The interest of the ESP teacher towards the phenomenon of IT etymology results from the possibility of using it, in a coherent integrated manner, in the English Language for Computer Science practical course, with multiple aims. They are listed below, in an order that is not that of priorities, as the goals should be interwoven.

Thus, the materials resulting from the teacher's and/or students' searches on etymological explanations of newly coined IT terms can be seen as rich sources, depending on the learners' language proficiency level, of General English and Technical English class input, more specifically, in our case, Computer Science connected terminology. Therefore, tasks focused on *lexical development*, *grammar* and *skills* can be designed to an equal extent, in function of the course specific objectives and the students' investigated needs.

A considerable advantage of such resources is that the materials are authentic, as they are extracted from an immediate reality which is part of the trainees' field of professional interest for both study and, further on, work.

Moreover, as IT terminology does not have chances to become as easily perishable or disused as most of the content in the CS field does, the tasks designed based on them should exploit this time-validity feature, in order to get the capacity of being preserved in the course as long as possible.

In terms of *student attitude formation*, there are certain aspects that are worth mentioning as clear *advantages* of the proposed approach.

Firstly, as some students still tend to use online means just for the *copy-paste* type of plagiarism, assigning them Internet search tasks on a given topic which is also of interest for them professionally (as *IT terminology etymological mechanisms* indeed is) can be conducive to developing their increased awareness of what a fair correct approach to studying and working in their field of activity really means.

Secondly, the approach can contribute to developing the students' capacity of *selecting useful information* as an important soft skill, not only for their language proficiency increase, but also as a *transferable* ability to other areas in their activity as engineers.

Then, it can be a manner of sensitizing the future IT specialists to *staying tuned with the dynamic changes in their specific domain*, becoming more familiar with the kind of creativity and, generally speaking, atmosphere specific to that working environment, where humor, team spirit, good communication are musts. Using English in their professional activity in an appropriate manner will necessarily require such traits - and it is never too early to train our students in this respect.

A last, but certainly equally important advantage, as it was rightfully remarked in the literature (Tedre, 2006), the development of *creativity* in science in

general and in CS/IT in particular, an element which results from a complex of personal features the university should give shape to in the students, whether it refers to the creative use of the English language or to the design of programming languages. The instructional process should therefore take into account factors such as: the socio-economic status, political/ideological beliefs, background, culture and motivation of the individuals.

The features listed above can be developed by means of tasks and chains of activities meant to raise the students' awareness of the:

- IT domain rapid evolution, technically and linguistically,
- humor and spontaneity components, intrinsically incorporated in the field (as can be seen even from the motto to this study, for instance, which emphasizes the significance of finding a memorable name for an IT product in terms of its marketability and general impact),
- (multi)cultural and historical perspectives required to members of the IT community,
- linguistic possibilities of using IT terminology appropriately inside and, by extension, outside the domain proper,
- need - or rather the expectation – for creativity as a permanent attitude in their IT connected jobs.

In attaining such goals, the teacher's role is of utmost importance, as they should take the 'hat' of the course designer, of the researcher, but equally of the IT field 'connoisseur ' to a quite large extent, in order to be able to process the input materials provided by the students. It is actually a two-way process of teaching/learning taking place between the teacher and the learners, who are frequently more familiar with the technical field than the language trainer.

A range of task types and patterns based on the IT terminology etymology input is briefly suggested here. It is necessary to mention that they can be amended in order to answer the needs of each concrete educational context, depending on the students' profile and of the course specific goals.

Thus, lexis development tasks can go beyond the words and word families specific to IT, towards the 'lay' use of the same terms. There are numerous examples of this kind: *booting*, *kindle*, *swing* and so on. Grammar development tasks based on the input on the topic discussed can be also inserted in the chain, in function of the students' level and of each particular lesson objectives.

In terms of skills, an important point is that the task type should fully match the input type; for instance, on the basis of Internet searches for the etymology of a certain IT term, if the students have identified explanations coming under the form of narratives, these should be exploited as input for reading/listening, while speaking/writing activities should come to the fore if they identified texts of the process presentation type, for instance.

For higher language levels, various input texts can be used in order to ask the learners to study stylistic issues, making them internalize different registers to be identified in IT texts, which considerably vary in degree of formality, for instance, as well as in other specific constructions, occurring in various text types, from scientific articles to forums, blogs or social networks. The trainees' capacity of

detecting – and, later on, generate - humor is not less important, as this is a field in which genuine wit is highly appreciated.

The IT community members' effort to create names for their tools-of-the-trade, which should resist in time and be accepted worldwide for their powerful impact, is mentioned quite frequently by various authors discussing Computer Science terms etymology (Hoff, 2011).

In most cases, emphasis of the people coining the new terms is placed on the intention to find the *'perfect'* word that should reflect both its technical description, but always in combination with a hint to something very *'catchy'*, as the case was for the two-meaning term *twitter,* explained from its etymological perspective by two different sources, with both emphasizing the value of the term double-meaning (http://www.wattpad.com/49758350-twitter and http://voices.yahoo.com/ why-twitter-called-twitter-6377844.html) as: *'a short burst of inconsequential information,'* *and 'chirps from birds'. And that's exactly what the product was.*

It is precisely this kind of approach that the tasks should develop in the students, thus preparing them to face the real demands of their IT-specific working place environment.

4. Outcomes and open conclusions

As this approach is currently being applied for the first time, its preliminary results only can be discussed at this stage. The data that are used originate from the students' project and homework materials, so we can conclude, based on their multitude and variety, that there was a vivid interest among the learners towards the topic.

Most of their remarks in solving the designed tasks have also been included in creating the above table of etymology types specific to IT. The process has been an enriching one for the teacher as well, as the tertiary education bachelor and master level engineering students are definitely more familiar with the technical aspects in IT and with the registers and other linguistic and stylistic aspects that are domain-specific than their ESP trainer.

Therefore, if consistently conceived, designed and applied, the IT terminology etymology focused activities in the ESP class may have good chances to gradually raise the students' awareness of the:
-    stable forming paradigms of IT words,
-    dynamic evolution of their professional domain at international level, permanently developing new etymological mechanisms for the specific terminology,
-    optimal manner of using English in their jobs,
-    need for creativity and imagination as attitude components required from a Computer Science engineer in general.

BIBLIOGRAPHICAL REFERENCES
Hoff,   Tim,   (2011).   *10   Origins   of   Common   Internet   Terms*, http://listverse.com/2011/08/06/10-origins-of-common-internet-terms/.
Howe, Denis, (1985). *The Free On-line Dictionary of Computing*, http://foldoc.org/.

Tedre, Matti, (2006). *The Development Of Computer Science. A Sociocultural Perspective*, University of Joensuu, Finland, ftp://cs.joensuu.fi/pub/Dissertations/tedre.pdf.

The Etymology of the Terminology of Computer Science, (2014). http://web.presby.edu/~phmeeker/classes/pc/CSC450/Notes/The%20Etymolo gy%20of%20the%20Terminology%20of%20Computer%20Science.doc.

The Merriam-Webster Dictionary Online, (2014). *Where do new words come from? How do you figure out their histories?*, http://www.merriam-webster.com/help/faq/etymology.htm.

Topală, Dragos, (2009). *Jargonul Informatic: Între Terminologie şi Expresia Colocvială*, în Analele Universităţii din Craiova, Seria Ştiinţe Filologice, Lingvistică", Anul XXXI, nr. 1-2/2009, p. 397-402, ISSN 1224-5712, www.fabula.org/actualites/article34902.php, 2009.

W3Techs, *Usage of content languages for websites*, (2014). http://w3techs.com/technologies/overview/content_language/all.

CREATIVITY IN GENERATING COMPUTER SCIENCE TERMINOLOGY – AN ETYMOLOGICAL STUDY FROM THE APPLIED LINGUISTICS PERSPECTIVE

Abstract: The aim of the study is to investigate certain aspects of human creativity manifested in generating terminology in English specific to the domain of Computer Science/IT, which is of interest for the specific UPB students' profile, with applications in the teaching of English for Specific Purposes – ESP. As this research has been the result of a close collaboration between the ESP teacher, focused in this case on IT and her students at bachelor level, respectively, the use of the data has then been channelled towards the creation of a range of didactic applications for the ESP course. The specific terminology is studied and grouped by categories for which optimal classification criteria are attempted at, and the goals for which the scientific research has been initiated is reflected in a series of applied linguistics based guidelines: creating tasks for internalizing the field terminology, innovative means of specialized lexis acquisition, as well as activities for the development of essential information (re)discovering skills - in a field unavoidably subjected to a much faster obsolescence process than in any other human knowledge areas, meant to be further transferred by the students into lifelong language autonomous study abilities. Moreover, the thus generated glossary can represent, we maintain, a rich source for the development of 'lay' English lexis – a direction which is also exploited by designing specific tasks. The whole process involving the teacher and the students can be conducive, by reflection and personalization of the action, towards the personal and professional optimization of the participants.

Key words: creativity, terminology, etymology, ESP, IT.