XVII International Congress of Linguists
Prague, Czech Republic,
July, 24-29, 2003

Semantic nets in the Net

Mafalda ANTUNES (mca@iltec.pt) and Susana CORREIA (smc@iltec.pt)
Instituto de Linguística Teórica e Computacional (ILTEC)
Lisbon, Portugal

0. Introduction

Internet's field favours the birth of new terminological units due to the dimension of its growth and impact in current days. This justifies a more in-depth inquiry on some of this domain's specific vocabulary phenomena.

Anyone can verify in the internet the existence of a group of terms belonging to only one semantic net[1] and denominating different areas of virtual reality. Terms such as *internet*, *net, web, virtual, on-line*, as well as *e-*, and *ciber*, share some semantic features, although they are not always interchangeable, even if they fitted into the same syntactic categories.

We intend to demonstrate that these terms, apparently synonyms, cannot be interchanged within all contexts and, therefore, reveal a type of graded synonymy.

One can hear, in Portuguese Language, more and more frequently, expressions such as *e-learning, e-banking*, or *dicionário on-line*. Why then doesn't anyone hear the expression *learning virtual, banking virtual*, or even *e-dicionário*, assuming the apparent synonymy of the three terms (*e-, on-line, virtual*)? How then can we describe internet's new terms use in Portuguese, specially focusing synonymy, since they are borrowed from another language?

We will take on Leibniz's classical assumption about synonymy, according to which 'two expressions are synonymous if the substitution of one for the other never changes the truth value of a sentence in which the substitution is made' (cf. Miller *et al.* 1990: 241), in order to demonstrate that, like Miller *et al.* declare, synonymy, once defined that way, is rare or does not exist at all.

---

[1] The same as *semantic field*, in this context. See section 3 for a definition of semantic field.

We will assume the hypothesis that "the synonymy is best thought of as one end of a continuum along which similarity of meaning can be graded." (Miller *et al.* 1990: 241). Cruse (1986) proposes a synonymy scale within terms, too, that we will hold on in the terms analysis. We will thus try to demonstrate the internet terms synonymy grade or hierarchy.

These terms were only recently attested in Portuguese language reference dictionaries. Therefore we will consider first the way internet's discourse neologisms[2] and loan words[3] (in this case, anglicisms) were composed in the linguistic system of Portuguese.

Next we will try to analyse the extent into which these terms are synonyms. We consider its inter-changeability within a paradigmatic axis, since these units are not completely interchangeable and have specific domains of use, besides being members of different grammatical categories (although belonging to the same semantic field).

1. Methodology

AiT (Associação de Informação Terminológica – Terminological Information Association) and ILTEC (Instituto de Linguística Teórica e Computacional – Theoretical and Computacional Linguistics Institute) have cooperated in the *Glossário Pan-Latino de Comércio Electrónico* (Electronic Commerce Glossary of all Latin Languages), a project of the Rede Pan-Latina de Terminologia (Realiter - Terminology Network of all Latin Languages), taking part in the settlement of the Portuguese equivalents of a terminological units' list. At the same time, it became evident the existence of a group of internet discourse terms that are frequent and productive in this field, the use of which we will try to describe.[4]

---

[2] Neologism, according to Alain Rey's work (1976), can briefly be defined as a lexical unit not attested in the active vocabulary of a language on the moment immediately before the considered one.

[3] According to the *Dicionário de Termos Linguísticos*, 2nd vol. (1990), a *loan word* is a "lexical borrowing not integrated in the language, being foreign in its phonemes, inflection, and even in its graphic form, or else a Portuguese word taken with the foreign words meaning used in much the same way" (our translation).

[4] This work follows one that exclusively dealt with the description and classification in Portuguese of terms in which occurred the particle *e-* (the same as *electronic*), such as *e-mail, e-commerce, comércio-e, e-pipoca* or *e-nigma*. These unities were called *e-terms*. Cf. ANTUNES, M., CORREIA S. e GONÇALVES, R. (2002). "E-termos: descrição e hipóteses de classificação", *in Actas do XVIII Encontro Nacional da APL*, (Outubro de 2002), Lisboa: APL e Colibri (in print).

First we selected the most productive terms, searching on the internet (search engines), and concluded that:

internet, net, web, ciber, on-line, virtual, e-

were the most frequent units in the internet discourse semantic field.

Secondly, it was built a list with their contexts and occurrences. In this list we tried to point out the differences and similarities, from the semantic point of view, between these terms belonging to internet's semantic field. In order to do this, we selected the most significant and frequent contexts and pointed out the occurrences (or absences) of all the terms in those contexts.

The analyzed terms were taken from (European and Brazilian) Portuguese sites and specialized publications. Note that we only took into consideration the (European and Brazilian) Portuguese occurrences. Actually, the analyzed terms are borrowed, but they are already inserted in Portuguese structures.

The most frequent contexts of occurrence were:

– *navegar* (em)
– *estar* (em)
– *nauta*
– *café*
– *business*
– *banking*
– *learning*
– *commerce*
– *biblioteca*
– *universidade*
– *escola*

For the definitions we followed the *Dicionário da Língua Portuguesa Contemporânea* (DLPC). In the semantic decomposition we selected the relevant semantic components, in order to emphasize the proper features.

2. Neology and the loan words integration – the case of the internet terms

The neological nature of the terms scrutinized here is responsible for some of the assistematicity we find in its adaptation to the Portuguese linguistic's system. Thus, although these are attested terms, they show up in different ways:

– *Internet*: *<Internet>*[5], *<internet>*: irrespectively of being written with capitals or small letters, it is always a noun with the feminine gender. Of all the terms considered here this one shows the most consistent graphic representation.

– *Net*: *<net>*, when isolated it has the feminine gender (eg.: a *Net*, a *Rede*), when agglutinated (eg.: *netnauta*, *netcafé*, *netescola*) its gender depends on the compound's second element.

– *Web*: behaves quite similarly to *net*.

Net and *web* show two kinds of behaviour: they are either autonomous or comopositional elements – e.g.: *vou consultar a web* (*I'm going to the web*), *vou aparecer no webcafé às 21h* (*I will show up in the webcafe at nine p.m.*).

– *Ciber*: *<cyber>*, *<ciber>*, written in the original form (anglo-saxonic form), or adapted to the Portuguese ortographic system (more frequently).

– *On-line*: *<on-line>*, *<online>*, written with or without hyphen (in both cases in the original form); it can also appear as the Portuguese translated form (*em linha*) as it is proposed in DLPC.

– *E-*: *<e-commerce>*, *<comércio-e>*, *<ecommerce>*, placed before or after a noun, with or without hyphen.

– *Virtual*: <virtualnauta>, <café virtual>, placed before or after a noun; always without hyphen (separated or agglutinated).

As it can easily be observed, these neologisms and loan words integration in the Portuguese structures is still not concluded and reflects the speakers'

---

[5] "< >" represents the graphic form in which these terms show up.

capacity to integrate in their language, in a productive and creative way, items that previously didn't belong to it. Lyons (1977: 69-70) differentiates the concept of 'productivity' from the concept of 'creativity', claiming the following: "whereas productivity is a characteristic inherent to the linguistic system, creativity is the speaker's ability to extend the linguistic system by means of unpredictable, though motivated, abstraction and comparison principles" (our translation). Productivity is, thus, spontaneous and unaware (according to the speaker's grammar), creativity is motivated, deliberate and conscious. In the case of Portuguese integration system of the internet terms, we found both phenomena: there are structures that follow the Portuguese grammar (such as, taking the Portuguese default order noun-adjective: *dicionário virtual*, *dicionário on-line*), and there are structures clearly motivated, for example, by phonetic games.

3. Semantic fields, denomination and synonymy

The selected terms, as it was said before, belong to the same semantic field. The *Dicionário de termos linguisticos* (2nd vol.) defines *semantic field*, taking Lyons (1977: 205), as "term that refers to a group of lexemes or other lexical units that are semantically bounded". (our translation)

In order to fill in internet's discourse multiple denominative[6] needs, that is, internet's discourse semantic field, numerous terms apparently synonymous were created.

In Cabré's view (1993: 213), mappings between denomination and concept are not univocal but multiple. So, it is possible for one single denomination to map onto different meanings (polissemy) and for one single concept to be denominated by different forms (synonymy). According to this author, even within specialized languages, whose different denominations are supposed to express one and only one different concept, synonymy appears in five ways: a) between a denomination and its definition (the dictionaries case), b) between a denomination and the same concept's illustration, c) between different languages equivalent terms, d) between different functional

---

[6] According to Correia (1999) "denomination" is the "relation established between a lexical unit and a mental image corresponding and representing a category of entities of the extralinguistic reality". (our translation)

languages denominations, e) between alternative denominations of the same historical language.

In agreement with this author, we consider that there is place for synonymy in specific domain languages exclusively when the terminological unities are equivalent in a formal and semantic way and belong to the same historical language.

This is what appears to be happening with some of the internet terms we inquired. It is worth noting that some of them are not equivalent, although there is the formal and semantical possibility of them being interswitched in certain contexts. We will try to show it in the next section.

4. Data observation

Now we will look at synonymy and questions related to it. All of the scrutinized terms are apparently interswitchable in certain contexts. In other contexts, however, this substitution is impossible. We consider that two words are strictly synonymous only when interswitchable in every context. We will thus try to present a synonymy hierarchy.

4.1. *Internet, Net* e *Web*

The term *Internet* is defined by DLPC as a noun that denominates the "international informatic network that allows one to receive and search for information by means of a computer". (our translation)
Net remits *to* Internet, which means that it is taken as a denominative variable of the same concept. Because of this, we are forced to believe that these terms share semantic components*: international network, receive and search for information.* Notwithstanding, it must be added that the term *net* has a specific morphological component*: Internet clipping.*

As we see it, *Internet* and *net* have the same semantic component, although one can occur as a composing element (*net*) and the other doesn't (*Internet*). That similarity only takes place because they are recently born and current discourse terms. They would probably be much more specialized if occurring in a more specialized discourse. They are, then, true synonyms

from a semantic features point of view, though not occurring in the same contexts because of formal reasons.

The reduced form is always much easier, specialy if we look at the context. Having its origin in *Internet*, and being a reduced form (monossylabic), *net*'s use makes it easier to behave like a composing element (*Netcabo*, netcable, *net-café*, net-café), which rarely occurs with the former.

The next table shows the semantic decomposition relevant components of the terms *Internet* and *Net*, making clear their semantic similarities and their formal difference.

Terms	Relevant semantic components		Formal aspect
	International system	Search/reception of information	
Internet	✓	✓	Extended form
Net	✓	✓	Reduced form

Figure 1 – semantic and formal decomposition of *Internet* and *Net*.

Once this is showed, it is easy to explain the occurrence of expressions like "Estou na (Inter)Net", and "Netnautas", *Netnautas* being "those who navigate on the Internet", and the absence of expressions like **Internetnauta*.

Web *is defined as a "system of access to the Internet information with the hypertext form" (DLPC). (our translation) The term* web *is slightly different from the previously presented: it is a "system" that allows "to access to Internet". By the data observation we see that* Net *and* Web *have a very similar behavior: they are both nouns used as a designation of the entering into the* system *or the* international network of information search*, in one hand; and they are both used, not only as common nouns (being on the* net/web*, navigating on the net/web) but also as composing elements, on the other. There are a lot of composed terms having with one of these elements (*net *or* web*) followed or preceded by autonomous lexical unities such as* café *(café),* escola *(school) or* nauta *(naut), because of the reduced form of those terms.* Europa *is clearly a similar case, being a noun that originated the name*

of the new monetary unit, the euro*, that in his turn originated terms such as* euromercados *(euromarkets) (trade markets of the* euro *area) or* eurotransacções *(eurotrades) (trades made in the* euro *area).*

The next text box presents a semantic decomposition scheme, based on the relevant components, that outlines the formal difference of the terms *internet* and *net*.

Terms	Relevant semantic componets			Formal aspect
	International system	Search/reception of information	Hypertext form[7]	
Internet	✓	✓		Extended form
Net	✓	✓		Reduced form
Web	✓	✓	✓	

Figure 2 – semantic and formal decomposition of the terms *Internet*, *Net*, and *Web.*

As a conclusion, we can now state that the term *net* shares semantic features with *internet* and that *web* shares them with *net* and *internet*. This is a case of specific lexical and semantic relations: inclusion and overlapping (Cruse, 1986: 86-87). Thus, we propose a synonymy hierarchy, represented in the following sheme:
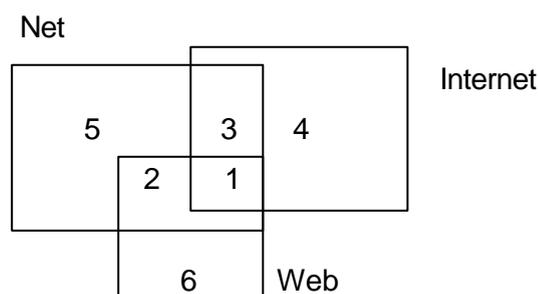

Figure 3 – ranked lexical relation between *internet*, *net*, and *web*.

Where:

1 – semantic features shared by *internet, net,* and *web.*
2 – features shared by *net,* and *web.*

3 – features shared by *net*, and *internet* – variants of the same term.

4 – peculiar to *Internet*.

5 – peculiar to *net*.

6 – peculiar to *web*.

Web and *net* share the possibility of becoming composing elements. *Web*, *net* and *internet* altogether are international systems of information access.

4.2. *On-line*, *virtual*, *e-*(electronic) and *cyber*

The scrutinized adjectives revealed peculiar properties too. In order to account for the similarities and the differences a table was built where the semantic components considered relevant to the synonymy hierarchy analysis are presented.

Thus, and according to DLPC, *on-line* remits to *em linha*, whose meaning is 'connected to the Internet'. Therefore, *Caixa Geral de Depósitos on-line* is the name for the banking institution site where it is accessible on the Internet. In his turn, there's nothing like **banking on-line* or **on-line banking* (sic). The notion of communication and Internet connection is inherent to the term *on-line*.

A form that is semantically very close to *on-line*, often behaving as its synonymous, is the adjective *virtual*. Its definition within the computational language is "that which is analogic", that is, that which is software simulated. *On-line* semantics gets, thus, another semantic component that we can identify as "simulation". *Realidade virtual*, for example, is precisely a simulated reality.

E- shares the semantic components above mentioned, but it has an extra formal component: it is a reduced form of *electronic /electrónico*. An *e-mail* is an electronic message and *e-business* is electronic business, that is, made in Internet.

Ciber- is a compositional element that expresses both notions of electronic communication and virtual reality. However, being a reduced form

---

[7] Costa e Melo (2002): "disposition of the segments of a text that allows related access" (the same as 'link')".

of *cybernetics*, it has another semantic feature: movement. We can see it in the examples: *ciber-espaço* (cyberspace) and *cibernauta* (cybernaut).

Next we present the analyzed terms semantic composition scheme:

Terms	Relevant semantic components				Formal aspect
	Communication	Connection to the Internet	Simulation	Movement idea	Compositional element
On-line	✓	✓			
Virtual	✓	✓	✓		
E-	✓	✓	✓		✓
Ciber	✓	✓	✓	✓	✓

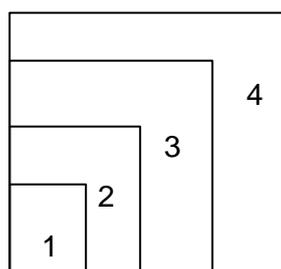Figure 4 – semantic and formal decomposition existing in *On-line*, *Virtual*, *E-*, and *Ciber*.

Figure 5 – lexical relation hierarchy set between *On-line*, *Virtual*, *E-*, and *Ciber*.

Where:

1 – *on-line* semantic features (shared by the other terms).

2 – *virtual* semantic features.

3 – *e-* semantic features.

4 – *ciber* semantic features.

We detect within this group of terms the lexical relations of semantic features inclusion and overlapping. From a semantic point of view, *ciber* is the most complex term, having a distinctive semantic component absent from the other terms.

The following summary-table of occurrences takes as a reference some of the most frequent contexts and points out the terms use differences within Internet discourse.

Occurrences of some *Internet* semantic field terms						
	Nouns			Adjectives			
	Internet	Net	Web	Cyber	On-line	Virtual	E-
Estar (na/o) ---	✓	✓	✓	✗	✓	✗	✓
Navegar (na/o) ---	✓	✓	✓	✗	✗	✗	e-navegar(?)
--- + nauta	✗	✓	✓	✓	✗	✓	✓
--- + café	✗	✓	✓	✓	✓	✓	✗
--- + business	✗	✗	✗	✗	✓	✗	✓
--- + banking	✗	✗	✗	✗	✗	✗	✓
--- + learning	✗	✗	✗	✗	✗	✗	✓
--- + commerce	✗	✗	✗	✗	Comércio on-line	Comércio virtual	✓
--- + biblioteca	✗	✗	✗	✗	✓	✓	Raro
--- + universidade	✗	✗	✗	✗	raro	✓	✗
--- + escola	✗	✓	✓	✗	✓	✓	✓

Figure 6 – summary-table of occurrences of the terms *Internet*, *Net*, *Web, Ciber, On-line, Virtual*, and *E*, belonging to Internet's semantic field.

5. Final remarks

With this work, we concluded that:
- Internet discourse, being recent, is quite permeable to lexical innovations birth;
- these borrowed new lexical items are productively integrated in the Portuguese linguistic system, originating new forms;
- being a new field, the Internet discourse is vulnerable to changes because of its semantic components lack of stability. It is possible, therefore, *net* and *Internet* to suffer the semantic specialization that will separate their semantic components.

To sum up, we can also say that:
- It is possible to determine different levels of synonymy based on semantic decomposition and context analysis.
- Internet terms, though belonging to the same semantic field and being semantically close, are not true synonyms.

11

We conceived this work as a contribution to the description of some of Internet discourse terms behaviour. We keep in mind, however, that it is a first approach, since this is a subject in development and asking for being developed.

6. Bibliography

ANTUNES M., S. CORREIA e R. GONÇALVES (2002). *E-termos: descrição e hipótese de classificação.* Lisboa. Instituto de Linguística Teórica e Computacional (ILTEC) e Associação de Informação Terminológica (AiT).

AITCHISON, J. (1988) *Words in the mind – An introduction to the Mental Lexicon.* New York: Basil Blackwell.

ARNTZ, R. & H. PICHT (1995) *Introducción a la Terminoligía.* Traducción del alemán, Amelia de Irazazábal *et al.* Fundación Germán Sánchez Ruipérez. Madrid: Ediciones Pirámide.

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

CABRÉ, M. T. (1993) *La Terminología. Teoría, metodología, aplicaciones.* Barcelona: Editorial Antártida / Empúries.

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

CANUTO, Marie-Berthe Vittoz. (2001) «Le Lexique des Technologies de L'Information et de la Communication et son Intégration en Français et en Italien», In: *Cahiers de Lexicologie.* Paris: Honoré Champion.

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

CORREIA, M. (1989). «Euro- um novo prefixo do português?». Trabalho realizado para o Centro Português de Terminologia (não publicado).

CORREIA, M. (1995). «O léxico na economia da língua». In: *Ciências da Informação,* Brasília, vol. 24, nr.3, pp. 299-307.

CORREIA, M. (1999). *A denominação das qualidades – contributos para a compreensão da estrutura do léxico português.* Dissertação de Doutoramento apresentada à Universidade de Lisboa (não publicada).

COSTA, J. Almeida & A. Sampaio e MELO (2002). Dicionário da Língua Portuguesa 2003. 9ª edição. Porto: Porto Editora.

CRUSE, D. A. (1986). *Lexical Semantics.* Cambridge: CUP.

KLEIBER, G. (1984). «Dénomination et relations dénominatives». In: *Langages*, nº76, pp. 77-94.

LYONS, John (1977). *Semântica–I.* Lisboa: Editorial Presença (trad. portuguesa de Wanda Ramos).

MILLER *et al.* (1990) «Introduction to WordNet: An On-line Lexical Database». In: *International Journal of Lexicography*, vol. 3, nr. 4. Oxford: OUP.

REY, Alain. (1976). «Néologisme: un pseudoconcept?» In: *Cahiers de Lexicologie,* n.º 28, pp. 3-17.

TEMMERMAN, R. (2000) *Towards New Ways of Terminology Description – The sociocognitive approach.* Amsterdam / Philadelphia, John Benjamins.

XAVIER, M. F. e M. H. M. Mateus (1990). *Dicionário de Termos Linguísticos*, vol. II. Lisboa: Edições Cosmos.