

# *In the same boat and other idiomatic seafaring expressions*

Rita Marinelli\*, Laura Cignoni\*

\*Istituto di Linguistica Computazionale, C.N.R.  
Area della Ricerca Via Moruzzi 1, 56124 Pisa, Italy  
e-mail: [Rita.Marinelli@ilc.cnr.it](mailto:Rita.Marinelli@ilc.cnr.it), [Laura.Cignoni@ilc.cnr.it](mailto:Laura.Cignoni@ilc.cnr.it)

## Abstract

This paper reports on a research carried out at the Institute for Computational Linguistics (ILC) on a set of idiomatic nautical expressions in Italian and English. A total of 200 Italian expressions were first selected and examined, using both monolingual and bilingual dictionaries, as well as specific lexicographical works dealing with the subject of idiomaticity, especially of the maritime type, and a similar undertaking was then conducted for the English expressions. We discuss the possibility of including both the Italian and English idiomatic expressions in the semantic database *Mariterm*, which contains terms belonging to the maritime domain. We describe the terminological database and the way in which the idiomatic expressions can be organised within the system, so that, similarly to the other synsets, they are connected to other concepts represented in the database, but at the same time continue to belong to a group of particular linguistic expressions. Furthermore, we study similarities and differences in meaning and usage of some idiomatic expressions in the two languages.

**Keywords:** Terminology, Semantic databases, Idiomatic expressions

## 1. Introduction

Italy and England have a rich maritime history and therefore it comes as no surprise that there are so many maritime words and expressions which originated on the sea, deriving from the customs and traditions of seamen who spent more time on ship than on land, and therefore carrying with them a large amount of cultural information.

Nowadays, in the global economy, the port is a focal point of the network which involves not only the commercial, but also industrial, economic, financial, and logistic activities.

The activities of the port have been extended beyond its geographical limits, interconnected as they are with the surrounding territory. In this respect, an ever increasing number of individuals are involved in the field of work (brokers and shipping agencies, freight forwarders, export-import companies, etc.), of school and didactics (nautical institutes, professional training, etc.), of maritime law (maritime and insurance contracts, averages and damages, passenger rights, etc.).

We were asked by specialized professional users working in the field of maritime navigation and transport to implement a terminological resource acting as a support for the definition and translation (Italian - English) of the terms belonging to this domain (Marinelli *et al.*, 2003).

Firstly, a large set of terms was codified containing semantic information about the knowledge field of navigation; secondly, a set of specialized concepts belonging to maritime transport terminology were also inserted in the database.

Changes caused by the globalization of trade, business and travel, together with the development of technology for communication, translation and dissemination of terms within the maritime activity and its related terminology were also taken into account.

The “world” of shipping and maritime transport, comprising

both passengers and goods, is to be considered a center of interest connected to other fields such as tourism or intercultural communication. Therefore, we decided to include in the database, mostly rich in technical terms, the idiomatic expressions and sayings originating from this domain, which is a cultural environment affecting - also from a lexical point of view - our everyday life. In this way, the lexical coverage of the database *Mariterm* has been increased with concepts connected to navigation and belonging to a cultural linguistic heritage shared by and in which so many countries identify themselves.

## 2. The Terminological Database

*Mariterm* is a database organized in semantic relations, following the EuroWordNet (EWN)/ItalWordNet (IWN) model, and according to the Princeton WordNet (WN) philosophy (Vossen, 1999; Miller *et al.*, 1990). The model used to structure this terminological lexicon is based on the concept of *synset*, a set of synonymous words, and on the vision of a synset defined by semantic relations which connect it to other synsets and reflect the user’s mental lexicon (Miller and Fellbaum, 2007).

The database contains semantic information for an approximate 3,500 lemmas belonging to the maritime knowledge field, and clustered into approximately 2,500 synsets, e.g.: *timone, governale* (helm); *naufragare, colare a picco, affondare* (to sink) (Marinelli and Roventini, 2006). The lexical semantic relations provided by the model can be described as: 1) relations linking the terms with other terms within the database (or ‘internal relations’); 2) relations connecting the terms to the synsets of the generic English WordNet (WN) (or ‘equivalence relations’).

### 2.1 Internal Relations

Synonymous meanings are grouped in a synset, and

language-internal relations hold between pairs of synsets. WNs are based on the management of conceptual structures hierarchically organized by means of ‘vertical’ relations of the type “has\_hyperonym/ has\_hyponym”, which provide coherence and consistency.

In the EWN/IWN linguistic model a variety of lexico-semantic relations (horizontal relations) such as ‘part\_of’, ‘cause’, ‘purpose’, ‘sub\_event’, ‘belong\_to\_class’, etc. are also used to represent the organization of lexical knowledge. Horizontal relations are employed to indicate semantic relations between concepts which are neither synonyms nor hierarchically-dependent terms (Marinelli *et al.* 2010):

*polizza di carico* (bill of lading) has \_hyperonym *polizza* (insurance)  
*attraccare* (to dock) involved\_agent *ormeggiatore* (mooring man)

## 2.2 Equivalence Relations

The design principles of the database allow for lexical semantic relations connecting the terms within the specialized lexicon, as well as semantic relations linking the synsets of the terminological database to the general WordNet lexicon.

According to the conceptual architecture of the model, equivalence relations link the terminological synsets to their closest equivalent concepts of the Princeton WN through the Inter Lingual Index (ILI). The following relations are possible:

a) synonymy (or near synonymy) relationship:

*alare* eq\_near synonym *haul, cart, drag* (draw slowly or heavily)  
*stivatore* eq\_synonym *stevedore, loader, dockhand* (a labourer who loads and unloads vessels in a port)

b) inclusion of the synsets in a taxonomic chain:

*porto di trasbordo* (hub port) eq\_has\_hyperonym *seaport, haven, harbour* (a place where ships can take on or discharge cargo)

c) ‘part’, ‘role’, ‘means’, etc. relations, which help define the semantic field of each term more precisely:

*stivatore* (stevedore) eq\_role *to stow, to load* (fill by packing tightly)

These links to the ILI make it possible to connect the terms to the Top Ontology (TO) (Vossen, 2004), which is an abstract, hierarchically-organized and language-independent set of concepts:

*navigazione* (navigation)→ Agentive, Dynamic, Purpose,  
*tassa di nolo* (freight rate)→ Dynamic, Possession, Quantity.

## 3. Idiomatic Expressions within Mariterm

The usage of a large number of maritime terms with an extension of meaning (metaphorical or metonymic) in linguistic contexts of everyday life was highlighted in preliminary studies (Marinelli, 2008a; Marinelli, 2009). Sets of nouns: *rotta* (course), verbs: *approdare* (to berth), and idiomatic expressions belonging to the maritime domain: *arrivare in porto* (to reach the port), *avere il sopravvento* (to get the upper hand), etc. were studied. The results of our lexicographic analysis were most of all aimed at considering

the relationship between the maritime domain as “source domain”, i.e. the domain supporting the literal meaning of the expression, and the concept described in the “target domain”, i.e. the domain the sentence is actually about (Croft and Cruse, 2004), within the framework of Conceptual Metaphor Theory (Lakoff, 1993; Grady *et al.*, 1999, Evans, 2006)).

The current goal of our project is not only to focus on idiomatic nautical expressions and sayings in Italian and English, but also on the methodology employed to encode these expressions in the database, adding if possible information of a historical and/or philological type, and trying to structure this set of expressions, so as to enlarge the semantic net and enrich the terminological resource.

The list of words and expressions was obtained analyzing the contexts of usage of terms included in the CLIC corpus<sup>1</sup>, a large textual corpus of contemporary Italian (Cignoni and Coffey, 1998; Marinelli, 2003; Marinelli, 2008b), and from the intersection of a number of dictionaries, web sources and specialized texts. The main problem has been to manage these words and expressions in order to insert them in the database, embedding them in the conceptual structure of the resource, connecting and even blending them with the other synsets.

Synsets are known to contain different word-forms with the same meaning, as well as word-forms with more than one meaning appearing in as many different synsets as they have meanings (Fellbaum, 1998, 2011; Osherson and Fellbaum, 2009).

For example, the string *mare* (sea) occurs in Mariterm as a noun in two different synsets, and therefore has two distinct senses. The former expresses the concept *mare1*, defined as *grande massa di acqua* (large body of water), the latter *abbondanza, grande quantità* (abundance, great quantity), indicating anything apparently limitless in quantity or volume as in the phrase “*un mare di guai*” (a sea of troubles), as in the examples from the CLIC Corpus:

*Burrasche e \*mare\* mosso hanno costretto al ritiro 56 imbarcazioni* (literal use)

*Onde altissime e \*mare\* gelato; una prova di forza dal fascino antico* (literal use)

*Il governo è invece naufragato in un \*mare\* di polemiche e incomprensioni* (extended use).

*Gli organizzatori hanno pensato di svecchiare l'evento proponendo un \*mare\* di iniziative rivolte ai ragazzini* (extended use).

*Come potete immaginare ci ha lasciato in un \*mare\* di guai mentre lui tranquillo prepara cocktails* (extended use, idiomatic expression).

In WN3.1, the most recent version of WordNet, two different senses of the synset “sea” are represented, but there is no particular relation codifying the extension of meaning:

S: (n) **sea** (a division of an ocean or a large body of salt water

<sup>1</sup> Aim of the LE PAROLE project (Marinelli *et al.*, 2003) was to create a comparable set of large written language resources for all European languages. The Corpus, later renamed CLIC, was enlarged so as to reach 100 million words, covering the years until 2005.

partially enclosed by land)

S: (n) ocean, sea (anything apparently limitless in quantity or volume).

In these and other similar cases, instead of treating *mare1* and *mare2* as different meanings of the same word-form, we decided to highlight the second sense of *mare* (sea) as an ‘extended’ sense of the former, using a particular relation which encodes the synset extensions.

A semantic relation was created and included in the set of database relations so as to codify the sense extension phenomenon in the case of metaphorical or metonymic uses of a term, namely the relation *has extension* and its reverse *is extension of*, already exploited to represent particular ‘extended’ senses of Proper Names (Cignoni and Coffey, 2003; Marinelli, 2010):

*mare1* (grande massa di acqua salata) *has hyperonym* acqua (water)

*mare2* (abbondanza, grande quantità) *has hyperonym* quantità (quantity)

*mare1 has extension mare2*

*mare2 is extension of mare1*

The inclusion of idiomatic expressions should occur in an organized manner so that, similarly to the other synsets, they can be connected to other concepts represented in the database, but at the same time continue to belong to a group of particular linguistic expressions.

Each idiomatic expression of the database is linked to a ‘thematic’ term, namely a keyword representing a central relevant concept in the specialized lexicon and occurring in associated idiomatic expressions. The semantic field (‘node of information’ connected to the given keyword) is supplied with new information which contributes to enriching the

entire conceptual network in which that term is not only included but also plays a central role. The link is constituted by a specific *ad hoc* relation connecting the idiomatic expressions to the relative thematic word(s), e.g.: *perdere la bussola* (→*bussola*), *tirare i remi in barca* (→*remo*, →*barca*). In the examples reported here we have decided to link the idiomatic expressions to the terms which are the actual root-words, the words related to the specific maritime language. Furthermore, these words remain fixed within the expression, and are considered together with the variants of the expression proper. In the example *tirare i remi in barca* (lit.: pull the oars in the boat) and its variants *appoggiare* (to rest), *mettere* (to place) *i remi* (the oars) *in barca* (in the boat), in other words *to ship the oars*, it is the concept, the mental representation of the oars in the boat that we assimilate to the staticity of the oars that we are pursuing; whether we are resting, pulling or placing them is irrelevant. All these kinds of data, information and semantic relations are managed by a tool which allows visualization and updating of the database. Each idiomatic expression is included in a set of items appropriately predisposed and managed by the tool, so that the idiomatic expressions form a repository of idioms that can be retrieved and listed on demand. Like the other synsets in the database, each expression is provided with a brief definition and is connected to its thematic term through a particular relation; if the term, for example *mare* (sea), appears in an idiomatic expression, the new relation *in idiomatic expression* and its reverse *idiomatic expression from* are used:

*mare1 in idiomatic expression* portare acqua al mare (fare una cosa inutile)

portare acqua al mare *idiomatic expression from* *mare1*

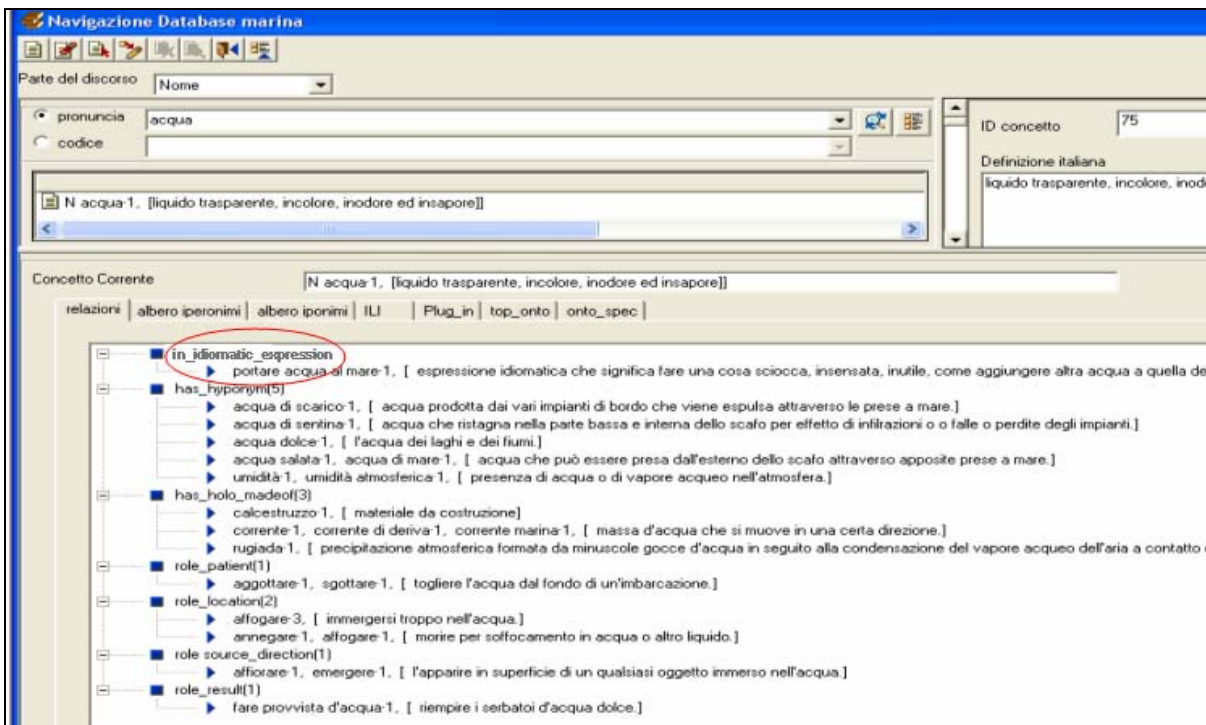


Fig. 1 Example of the term ‘acqua’

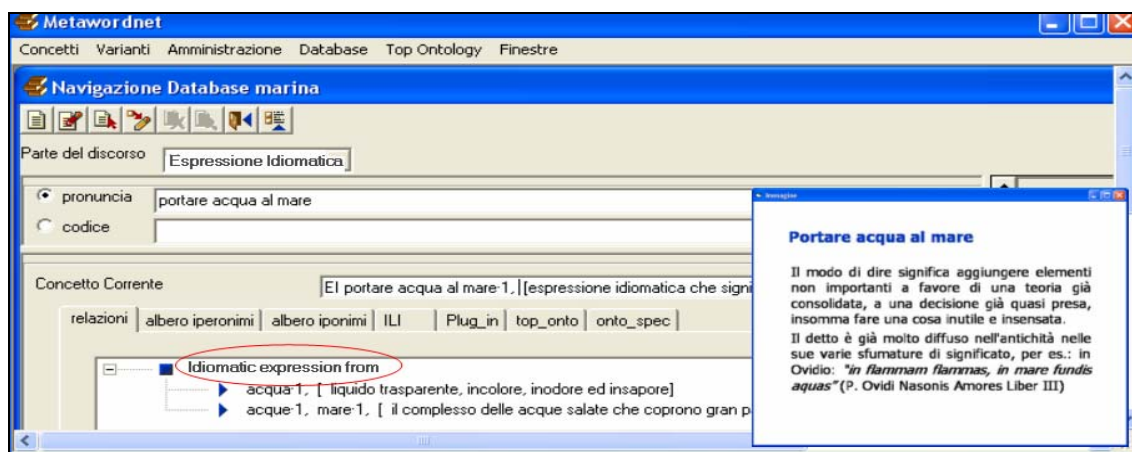


Fig. 2 “Portare acqua al mare”

In some cases, the tool displays a slot containing historical information concerning the idiomatic expression sought for, as in figures 1 and 2 above.

#### 4. English Translation

The idiomatic expressions included in the database also have the English translation, as provided by the conceptual model of the database.

When compared with nautical English phrases, many similarities were found to exist in the two languages, but sometimes a nautical expression was idiomatic in one language but not in the other (Cignoni *et al.*, 1999).

As far as the English translation is concerned, different situations may occur:

- a) The idiomatic Italian expression has an English translation which expresses the same concept, and uses terms belonging to the same maritime lexicon. Some expressions have a literal translation as well as an idiomatic sense which are identical, i.e. *un pesce fuor d'acqua* (a fish out of water), the same in the two languages, indicating unease and discomfort, or *parlare come uno scaricatore di porto* which is similar to “to talk like a dockworker”, or *buttato ai quattro venti*, which is identical to “scattered to the four winds”: if something is scattered to the four winds, it moves in all directions.
- b) The idiomatic expression has an English translation expressing the same concept, but using terms that do not belong to the same (maritime) ‘world’ i.e. *essere un porto di mare*, finding oneself in a very popular and crowded place (to be like Piccadilly Circus), *un mare di guai* (a heap of trouble), *goccia nel mare* (drop in the bucket), the last expression meaning that a certain amount of something is insufficient.
- c) There are some English idiomatic expressions belonging to the maritime ‘world’, with an equivalent expression in Italian; in other words they have the same meaning, but do not belong to the maritime ‘world’, i.e. *bruciare i ponti*, (lit.: to burn the bridges) which is equivalent to the English “to

burn one’s boats”, in the sense of leaving oneself with no means of escape; *buttarsi a terra*, - where the Italian uses the very generic “throw yourself to the ground”-, while the English uses “hit the deck”, the part of the ship, with the intention or meaning of protecting oneself. Another example, “sail close to the wind” has the same meaning of *camminare sul filo del rasoio* (to walk on a razor(’s) edge): if you sail close to the wind, you take the risk of doing something, of going close to the limit of what is allowed or acceptable.

#### 5. Conclusion

This paper has discussed the importance of Italian and English seafaring phraseology and the integration of idiomatic expressions in the terminological database Mariterm organized in semantic relations.

We related on i) the methodology designed to enhance the database with this set of expressions, in keeping with the conceptual model of the database; ii) the enrichment of the lexical resource with explanations of the origins of the most common expressions; iii) the similarities and differences in meaning and usage of some idiomatic expressions in the two languages, also using Italian and English corpora and databases as references.

We are planning to add data on the frequency in the corpus CLIC of a subset of idiomatic expressions. In this respect, it would be useful to create a link allowing visualization of a number of examples, together with an annotation regarding the use (either literal or extended) of the idiomatic expression considered in that particular context.

Corpora can help determine or at least suggest the canonical form of the idiomatic expression, on the basis of the frequency of the variants and analysis of the co-occurrences.

The relatively small number of expressions investigated in both languages will need to be expanded, and further studies will allow us to enrich the databases and provide a clearer picture of the phenomenon.

We suggest the possibility of using a database of this type - managed by a tool which provides an appealing visualization of the terms and their connections to other terms - also for didactic purposes, in both public and private schools, as well as for contexts in which the knowledge of nautical expressions is essential, for example for workers in yachting harbours and cruise ports, who need specific language when brought into contact with foreign tourists or peers from other countries.

With the present enhancement the resource is now enriched with different elements which can be considered part of a more general “culture of the sea”, rich in historical references, and based on roots which we have found to be international and without boundaries.

## 6. References

- Cignoni L., Coffey S. (1998). A Corpus-based study of Italian idiomatic phrases: from citation forms to ‘real-life’ occurrences, In *Proceedings of Euralex '98*, University of Liège, 1: 291-300.
- Cignoni L., Coffey S., Moon R. (1999). “Languages in Contrast”, *Idiom Variation in Italian and English, Two corpus-based studies*, II, 2, John Benjamins Publishing Company, 279-300.
- Cignoni L., Coffey S. (2003). At the interface of onomastics and phraseology. Multiword units as proper names - proper names as ‘common’ phrasal units, *Linguistica Computazionale*, XVI-XIX, IEPI. 1: 251-262.
- Croft W., Cruse D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Evans V., Green M. (2006). *Cognitive Linguistics. An Introduction*. Edinburgh, Edinburgh University Press.
- Fellbaum C. (1998). Towards a representation of idioms in WordNet. In *Proceedings of the workshop on the Use of WordNet in Natural Language Processing Systems*, Coling-ACL, Montreal.
- Fellbaum C. (2011). Idioms and collocations. In C. Maienborn, K. von Heusinger, P. Portner (Eds.), *Semantics*, De Gruyter Mouton, 2011. pp.: 441-456.
- Grady I., Oakley T., Coulson S. (1999). Blending and Metaphor. In *Metaphor in Cognitive Linguistics*, edited by G. Steen and R. Gibbs. Current Issues in Linguistic Theory 175. Amsterdam: John Benjamins.
- Lakoff G. (1993). “The Contemporary Theory of Metaphor”, *Metaphor and Thought*. Ed. A. Ortony (2nd edition). Cambridge MA: Cambridge University Press. 202-251.
- Marinelli R., Biagini L., Bindi R., Goggi S., Monachini M., Orsolini P., Picchi E., Rossi S., Calzolari N., Zampolli A. (2003). The Italian PAROLE corpus: an overview. In *Linguistica Computazionale*. XVI-XVII. IEPI. 1: 401-421.
- Marinelli R., Roventini A., Spadoni G. (2003). Linking a subset of maritime terminology to the Italian WordNet. In: *Proceedings of the Third International Conference on Maritime Terminology Communication and Globalization*. Lisbon. 2003. Lisboa, iLTeC (Instituto de Linguística Teórica e Computacional).
- Marinelli R., Roventini A. (2006). The Italian Maritime Lexicon and the ItalWordNet Semantic Database, In E. Miyares Bermúdez and L. Ruiz Miyares (Eds.), *Linguistics in the Twenty First Century*, Cambridge Scholar Press, 173-182.
- Marinelli R. (2008a). “Idiomatic Expressions and Metaphors from the Maritime Domain”. In M. Alvarez de la Granja. *Lenguaje figurado y motivacion. Una perspectiva desde la fraseologia*. Studien zur romanischen Sprachwissenschaft und Interculturellen Kommunikation. Frankfurt am Main: Peter Lang Internationaler Verlag der Wissenschaften. 209-220.
- Marinelli R., (2008b). Analisi di metafore e espressioni idiomatiche per mezzo di risorse computazionali e corpora elettronici. In *Critica del testo. L’Europa dei proverbi*. A cura di A. Punzi e I. Tomassetti. Roma. XI, nn. 1-2, 2008. 469-488.
- Marinelli R. (2009). Ontological Structure and Digital Corpora for Metaphorical Sense Recognition. In C. M. Bretones Callejas et al. (Eds.) *Applied Linguistics Now: Understanding Language and Mind / La Lingüística Aplicada actual: Comprender el lenguaje y la mente*, AESLA, Universidad de Almeria, Spain. pp. 1369-1380.
- Marinelli R. (2010). Ontological variation and sense extension in proper names. In *Proceedings of the XXVIII AESLA Conference, “Analizar datos > Describir variacion – Analysing data > Describing variation”*, Vigo, University of Vigo.
- Marinelli R., Mazzocchi F., Tiberi M., Motta M. (2010). Il modello semantico di EuroWordNet come strumento per la strutturazione della relazione associativa nei thesauri. *Bollettino AIB*, L, 3: 249-263.
- McCarthy M., O’Dell F. (2002). *English Idioms in Use*. Dubai: Cambridge University Press.
- Miller G. A., Beckwith R., Fellbaum C., Gross D., Miller K.J. (1990). Introduction to WordNet: An On-line Lexical Database, *International Journal of Lexicography*, III, 4: 235-244.
- Miller G. A., Fellbaum C. (2007). WordNet then and now, *Language Resources & Evaluation*. 41: 209-214.
- Osherson A., Fellbaum C. (2009). The representation of idioms in WordNet. In *Proceedings of ACL- IJCNLP 09*.
- Vossen, P. (1999). EuroWordNet General Document, 1999. <http://www.hum.uva.nl/~EWN>.
- Vossen P. (2004). Eurowordnet: a multilingual database of autonomous and language-specific wordnets connected via an Inter-LingualIndex. *International Journal of Lexicography*. 2004. 17 (2): 161-173.